THE TOPOLOGY OF TRANSFORMATION GROUPS*

P. A. SMITH

1. Stated in general terms, the problem which I wish to consider is the following: A group G of transformations operates in a space S; what relations must exist between the topology of this situation on the one hand, and its group theoretical properties, on the other? This is of course a rather vague question and I shall not attempt to describe all the recent results which could be considered as being relevant. I hope rather to illustrate certain special phases of the problem by means of examples. I shall then consider in some detail one special case where certain conclusions can be drawn which may possibly be of interest in algebraic geometry. I refer particularly to a generalization to higher dimensions of Harnack's theorem concerning the number of real branches which a real algebraic curve may possess.

Continuous groups

2. Let us begin by supposing that G is a continuous r-parameter group. The elements of G may then be thought of as points of a space which has locally the character of a euclidean r-space. Products and quotients of two elements are to vary continuously with those elements. What restrictions does the fact that G is a continuous group place upon the *space* G? It is very easy to see that G cannot be of arbitrary topological structure. For suppose that a is a fixed element different from the identity, and that x is an arbitrary element; then the transformation T_a: $x \to ax$, is a homeomorphism of G with itself, and no x remains fixed. Now if we assume that G is connected, a can be made to describe a continuous path toward the identity element, and T_a then undergoes a continuous modification into the identical transformation. That is, T_a belongs to the "class of the identity." But in many spaces (a sphere for example) every transformation which belongs to the class of the identity must leave at least one point fixed. Thus no two-parameter group can be topologically equivalent to a sphere.† It is just as easy to show that G must be

* An address delivered before the New York meeting of the Society on February 26, 1938, by invitation of the Program Committee.

† A necessary condition that a compact manifold admit transformations of the class of the identity without fixed points is that its Euler-Poincaré characteristic vanish. See Lefschetz, *Topology*, American Mathematical Society Colloquium Publications, vol. 12, New York, 1930, pp. 272, 359.

orientable (see for example [31]). It follows readily that the only surfaces which could be group-spaces are the torus, the infinite cylinder, and the euclidean plane; continuous groups of each of these types do exist.

Let us return to the situation mentioned in the beginning in which G is realized by means of continuous transformations in S. If we assume that G operates in S transitively, then S, as well as G, is restricted topologically. For example, if S is a surface, it must belong to one of six distinct topological types, at least if G is a Lie group, that is, analytic (Cartan [10]). In this special case of an analytic G, the topology of (G, S) is intimately related to the group theoretic infinitesimal structure of G. When G is compact this relationship reveals itself through a consideration of the invariant integrals on S (Cartan [8]). In particular, one may assume that S is identical with G, the transformations of the realization then being the T_a's defined above. The theory of invariant integrals then leads to an explicit enumeration of the Betti numbers of compact continuous groups. Cartan [9] has pointed out the suggestive fact that on the basis of these and related results, one may conclude that so far as the ordinary invariants of homology theory are concerned, every simply connected simple compact Lie group G is equivalent to the topological product of a finite number of spheres of varying dimensions. Whether or not G is actually homeomorphic to such a product remains to be determined.*

3. Let us now consider the problem stated at the beginning, from a somewhat different point of view. Given certain topological condi-

* Our present knowledge of the topology of Lie groups is due largely to the analytic and algebraic investigations of Cartan and Weyl. An excellent review of the situation will be found in Cartan's expository article [9]. Suffice it to mention here that the Betti numbers of compact groups were first obtained by Pontrjagin [22] by direct and elementary methods and subsequently by Brauer [5] by a method based on the theory of invariant integrals. The connecting link between the integrals and the homology groups was furnished by de Rham [24]. With regard to the realizations of G, it was pointed out by Cartan [10] that the study of the topological and group-theoretical properties of a Lie group G contains within itself the study of the properties of its transitive realizations. For suppose $\mathfrak{g}$ is a closed subgroup of G; then the cosets $x\mathfrak{g}$ (x in G) constitute, upon the introduction of natural topology, a space S such that the transformations $x\mathfrak{g} \to a(x\mathfrak{g})$ (a in G) define a transitive realization (G, S). Every transitive (G, S) is in fact equivalent to one defined in this manner by means of a suitably chosen $\mathfrak{g}$. The details justifying these remarks when the spaces G and S are quite general and analyticity is not assumed were first worked out by Freudenthal [12] and subsequently by the author in his study [33] of partially defined groups and realizations (the sort actually encountered in the classical theory of continuous groups). As concerns the homotopy groups of a continuous (non-analytic) G see Hurewicz [13]; for the special case of the Poincaré group, see also [33], [34], [35].

tions on (G, S), to what extent is the group-theoretic structure determined? Suppose, for example, one considers only the analytic case and asks: Given the dimension of S (a topological property), what are the continuous groups which can operate transitively in S? The answer constitutes, as we know, one of the important chapters in the classical works of Lie. Suppose however that one drops the assumption of analyticity and assumes merely that G and S are locally euclidean and that G, as well as the transformations by which G is realized, is continuous. Does one then have a situation which is essentially more general than that studied by Lie, or is every (G, S) of this type equivalent to a classical (analytic) one? In two rather difficult papers, Brouwer [6] showed that the continuous case is not more general than the analytic one when S is one-dimensional, and went far toward establishing a similar result when S is two-dimensional. With regard to G itself, we have now the theorem of von Neumann [18], somewhat sharpened by Pontrjagin [21], that every r-parameter G which, considered as a space, is compact, is topologically equivalent to an analytic G. Whether or not this holds for non-compact G's appears to be a difficult question.

These results have a special significance in connection with certain problems in the foundations of geometry which were occupying mathematicians of another generation. Helmholtz and Lie attempted to found euclidean n-dimensional geometry by assuming that space was an n-fold number manifold and then introducing certain postulates which were to define the group of rigid motions. It is well known that Hilbert criticized this procedure on the grounds that analyticity, postulated for the transformations which were to be rigid motions, did not properly belong to "foundations." Hilbert himself then proposed a set of axioms for two-dimensional geometry in which analyticity is not assumed, the decisive axiom being a topological closure requirement for the system of transformations (see Hilbert's *Grundlagen der Geometrie*, seventh edition). In this manner, the number plane received a truly topological conversion into the place of euclidean geometry. So far as I know, there has been little attempt to carry this development out in spaces of higher dimensions. The discovery that compact r-parameter groups are analytic is of course closely connected with this question; but there still remains the problem of characterizing, among the realizations (G, S) operating in a given number space S, those whose transformations, in their totality, are topologically equivalent to the rigid motions. A recent work of Montgomery and Zippin [16] which contains a topological-group-theoretical characterization of the group of rotations about a fixed

axis can, I believe, be considered as marking a certain progress toward a solution of the problem for three dimensions.*

FINITE GROUPS

4. I shall pass over the case in which G is an infinite discrete group and assume that G is finite. One may ask, what are the finite groups of auto-homeomorphisms of a space S with given topological properties? If S is homeomorphic to a sphere, the answer is known. The groups are essentially the same as the finite groups of orthogonal transformations of a sphere into itself; for example, if it is required that orientation be preserved, the possible groups are cyclic groups of rotations about a fixed axis and the rotation groups of the regular solids. This is a topological theorem and can be proved by topological methods (Kerékjártó [14]). If one asks the more special question, what are the finite groups of linear fractional substitutions in one complex variable, one may arrive at the answer (essentially the same as that of the preceding questions) by function-theoretic methods (Klein, *Lectures on the Icosahedron*). These methods, in their use of Riemann surfaces, have a certain topological element, and it is perhaps this fact that led Poincaré to express the belief that the determination of the finite groups of linear transformations in n variables must depend on the solution of topological problems.

5. I wish now to consider in more detail the special case in which G is finite and cyclic; that is, the case in which we are dealing with a single transformation t of finite period. Such transformations arise in a number of ways. Suppose, for example, that t is a birational transformation of an algebraic curve C into itself. Then it is known that if the genus of C is greater than one, t must be of finite period (the Schwarz-Klein theorem†). Again, in the theory of algebraic varieties, a particular type of involution, so-called, is that generated by a periodic transformation.‡ In certain problems in the calculus of variations one encounters cyclic and symmetric products. The m-fold topological product of a space K is the space S whose "points" are the ordered sets $(p_1, \cdots, p_m)$, (p_i in K). The m-fold symmetric product

* I am informed by one of these authors that they have recently established the still more significant theorem that every (G, S) in which G is compact and connected and S is a three-dimensional number space, and which possesses no identical transformations save that which corresponds to the identity element of G, is equivalent either to the group of rotations about a fixed axis or about a fixed point. Concerning characterizations of translation groups, see [17].

† See Severi-Loeffler, *Vorlesungen über Algebraische Geometrie*, p. 143.

‡ See, for example, Lucien Godeaux, *Les Involutions Cycliques Appartenant à une Surface Algébrique*, Actualités Scientifiques et Industrielles, no. 270, Paris, 1935.

$K^{(m)}$ of K is obtained from S by regarding two sets as identical if one is a cyclic permutation of the other. In determining the properties of $K^{(m)}$ it is of advantage to study the periodic transformation $t:(p_1, p_2, \cdots, p_m)\to(p_2, \cdots, p_m, p_1)$ of S into itself.*

The only invariant of G is now its order m, and our general problem therefore reduces to this: Does the topological nature of S impose restrictions on m? Suppose that S is an orientable surface of genus $p+1$ and that t preserves orientation. Then we have the interesting fact that m can not exceed $4p+2$.† I wish now to establish a similar result for higher dimensions. I shall suppose that S is an analytic orientable manifold of an even number of dimensions; S could, for example, be an algebraic surface without singularities; it would then have four (real) dimensions. I shall suppose also that t is analytic and preserves orientation. We shall see later that the invariant points of t, if any exist, constitute a finite number of submanifolds, each of an even number of dimensions. In a sense, we may regard as the general case that in which the invariant set is of dimension zero.‡ In any case, we shall assume that each of the transformations $t, t^2, \cdots, t^{m-1}$ leaves fixed only a finite number of points; when m is prime this condition holds automatically for the powers of t if it holds for t.

THEOREM. *If the conditions stated above hold, and if the Euler-Poincaré characteristic κ of S is negative, then there is an upper bound for the period m, which depends only on the Betti numbers of S.*

PROOF. To every isolated fixed point of a given transformation there can be associated a topologically significant numerical "index." Let O be a fixed point of t. In terms of suitable parameters at O, t can be represented by

$$u_i' = tu_i = \phi_i(u_1, \cdots, u_{2n}), \qquad i = 1, \cdots, 2n,$$

where the ϕ's are analytic and $\phi_i(0)=0$. If $J(u)$ is the Jacobian ma-

* Cyclic products $K^{(m)}$ were defined by Walker who determined their Betti numbers in terms of those of K. The case $m=2$ (symmetric products) had already been studied by the author [32] and by Richardson [26]. For applications in the calculus of variations, it seems to be necessary to know the Betti numbers of K modulo L, where L is a certain subcomplex, homeomorphic to K. These have been obtained by Richardson [27] for $m=2$ and by Richardson and Smith [29] for arbitrary prime m.

† First proved by Wiman [40] for birational transformations of algebraic curves; the first topological proof is by Steiger [37]. Nielsen [19] subsequently showed that the theorem holds if it is assumed that t^m is not necessarily the identity but merely belongs to the class of the identity.

‡ The fixed points of a projective transformation, for example, are isolated unless the characteristic equation has multiple roots.

trix of this transformation and I is the identity matrix, then it can be shown that the index at O is $+1$ if the determinant of $J(0)-I$ is greater than zero (Alexandroff and Hopf, *Topologie*, p. 537).* But in the present case this condition is bound to be satisfied, as the following considerations will show.

The space of differentials at O is transformed into itself linearly; and since this transformation is periodic, it leaves invariant a positive definite invariant form; in fact if we write $du_i=\xi_i$, such a form is

$$\sum_{i=1}^{2n} (\xi_i^2 + (t\xi_i)^2 + \cdots + (t^{m-1}\xi_i)^2).$$

On making a suitable linear transformation of parameters at O, the invariant form will become euclidean. Therefore, in the power series expansions of the functions ϕ_i, the coefficients of the linear terms may be assumed to form a proper† orthogonal matrix; hence, after a further linear change of parameters, t has, in the neighborhood of O, the form

$$\begin{aligned} tx_i &= x_i \cos\theta_i - x_{i+1}\sin\theta_i + X_i, \\ tx_{i+1} &= x_i \sin\theta_i + x_{i+1}\cos\theta_i + X_{i+1}, \qquad i = 1, 3, \cdots, 2n-1, \end{aligned} \tag{1}$$

where the X's are power series beginning with terms of at least the second degree, and the θ_i are multiples of $2\pi/m$. I assert that each θ_i is different from zero (mod 2π). Suppose for example that $\theta_1=0$. Then

$$\begin{aligned} tx_1 &= x_1 + X_1(x) = \psi_1(x), \\ tx_2 &= x_2 + X_2(x) = \psi_2(x), \\ &\cdots\cdots\cdots\cdots. \end{aligned}$$

Let

$$\Omega_i(x) = \psi_i(x) + \psi_i(tx) + \cdots + \psi_i(t^{m-1}x), \qquad i = 1, 2,$$

where, of course, tx means $(tx_1, \cdots, tx_{2n})$. Obviously $\Omega_i(x)=\Omega_i(tx)$. Moreover, it is easy to see that

$$\begin{aligned} \Omega_1(x) &= mx_1 + Y_1(x), \\ \Omega_2(x) &= mx_2 + Y_2(x), \end{aligned}$$

where Y_1, Y_2 begin with terms of at least second degree. Hence the transformation

$$\bar{x}_1 = \Omega_1(x),\ \bar{x}_2 = \Omega_2(x),\ \bar{x}_3 = x_3, \cdots,\ \bar{x}_{2n} = x_{2n}$$

---

* The condition is obviously independent of the choice of parameters at O.

† Proper, because orientation is preserved.

is (1, 1) in the neighborhood of O. We now examine the form which t takes in the variables $\bar{x}$. We have, for $i=1, 2$,

$$t\bar{x}_i = \Omega_i(tx) = \Omega_i(x) = \bar{x}_i.$$

Therefore t leaves fixed the points of the $(\bar{x}_1, \bar{x}_2)$-plane. This contradicts the assumption that the fixed points are finite in number and proves our assertion that no θ_i vanishes. Hence if we form the matrix $J(x)-I$ from equations (1), it follows immediately that its determinant is greater than zero for $(x)=(0)$. Hence O is of index $+1$.

Consider now the h-dimensional homology group H_h. It possesses a decomposition of the form

$$H_h = [\gamma_h^1] + \cdots + [\gamma_h^{R_h}] + Z_h,$$

where $[\gamma_h^i]$ is an infinite cyclic group generated by the cycle γ_h^i and Z_h is a finite group; R_h is the Betti number. The effect of t on the γ's is given by

$$t\gamma_h^i \sim \sum_j a_h^{ij}\gamma_h^j \pmod{Z_h}. \tag{2}$$

Let $\theta=\sum(-1)^h$ trace a_h, $a_h=\|a_h^{ij}\|$. In the Lefschetz theory of transformations the number θ is an important character of t, and when the fixed points are isolated, θ equals the algebraic sum of their indices (Lefschetz, *Topology*, p. 272). In the case under consideration it follows that $\theta \geqq 0$.

Now each matrix a_h is of period m; that is, its mth power is the identity. Consequently, the characteristic roots of a_h are mth roots of unity. It cannot happen that all these roots are $+1$, for then the trace of each a_h would equal R_h and θ would equal κ, which is impossible since $\kappa<0$. In exactly the same way we can say that not all the roots of all the matrices for t^d, $(1 \leqq d \leqq m-1)$, can equal $+1$.

The completion of the proof of the theorem now depends on two elementary theorems in number theory. Let $m=p_1^{a_1}, \cdots, p_s^{a_s}$ be the factorization of m into powers of distinct primes. Let $\Lambda(m)=p_1^{a_1}+\cdots+p_s^{a_s}$. Then first, if $\mu_1, \cdots, \mu_s$ is any set of integers whose l.c.m. is m, we have $\sum\mu_i \geqq \Lambda(m)$. Secondly, if R is any positive integer, the number of integers m for which $\Lambda(m) \leqq R$ is finite; we shall denote the largest one by $\chi(R)$.

Consider the "roots of t"; that is, the totality of characteristic roots of $a_0, a_1, \cdots, a_n$. Let the orders of these roots be $\lambda_1, \lambda_2, \cdots, \lambda_s, \cdots$, where $\lambda_1, \cdots, \lambda_s$ are a maximal set of *distinct* orders. Let l be the l.c.m. of $\lambda_1, \cdots, \lambda_s$. Then $l=m$. For, each λ_i is a divisor of m, therefore so is l. Suppose $l<m$; then the roots of t^l, being lth powers of

the corresponding roots of t, are all 1, which is impossible. From the preceding paragraph it follows that $\sum_1^s \lambda_i \geqq \Lambda(m)$. Let ϵ_i be a root of order λ_i, $(i=1, \cdots, s)$; it satisfies an irreducible equation $P_i(\epsilon)=0$ of degree λ_i with integer coefficients. P_i must be a factor of one of the characteristic equations. Since $\lambda_1, \cdots, \lambda_s$ are distinct, so are $P_1, \cdots, P_s$; therefore the sum of their degrees is less than or equal to the sum of the degrees of the characteristic equations; that is, $\sum_1^s \lambda_i \leqq \sum_1^s R_i = R$, say. Hence $\Lambda(m) \leqq R$ and $m \leqq \chi(R)$, which proves the theorem.

Whether or not the theorem remains true, if we drop analyticity and allow continua of fixed points, remains to be determined. It seems very unlikely that analyticity is really essential in the argument.

TYPE INVARIANTS OF PERIODIC TRANSFORMATIONS

5. Let us now adopt a slightly different point of view. Two homeomorphisms t and t_0 of S into itself are topologically equivalent, or belong to the same topological type if there exists a third homeomorphism τ such that $t_0=\tau^{-1}t\tau$. For arbitrary homeomorphisms, the problem of classifying types is too general to be of great interest, although a number of characteristic properties of types can be described in terms of the recurrence phenomena of repeated iteration (see for example [4]). But if t is periodic, one has invariants of a mere arithmetic nature, the period m for example. Suppose S is a circle. Then every periodic t of S into itself is equivalent to an orthogonal transformation.* In particular, if t preserves orientation it is equivalent to a rotation through an angle $2k\pi/m$, $(0<k\leqq m-1)$. If t_0 (of S into itself) is of the same type as t, then it is readily seen that m_0 must equal m, and k_0 must equal $\pm k$. Thus rotations of a circle through $2\pi/5$ and $4\pi/5$ do not belong to the same topological type. But suppose S is a torus with angular coordinates θ, ϕ and that t, t_0 are the rotations $\theta'=\theta+2\pi/5$, $\theta'=\theta+4\pi/5$, respectively. Then t and t_0 do belong to the same type (cf. Nielsen [20]). One can in fact see readily that if τ is the homeomorphism

$$\theta_0 = 2\theta + \phi,$$
$$\phi_0 = 5\theta + 3\phi$$

of S into itself, then $t_0=\tau t\tau^{-1}$. This suggests that the characteristic type invariants of periodic surface transformations will come to light only through a fairly penetrating analysis of surface topology. Such an analysis has recently been carried out by Nielsen [20] for orienta-

* First proved by J. F. Ritt [30].

ble surfaces and transformations which preserve orientation; the invariants obtained are sufficient to settle the question of whether or not two given periodic transformations are of the same type.

In Nielsen's work, the fundamental (Poincaré) group plays an important part. I shall now consider certain invariants which can be defined purely in terms of boundary relations. These invariants are less sharp than those of Nielsen, but they have proved to be very useful in a number of connections and can be defined for spaces of any number of dimensions.

Let $\mathfrak{g}$ be a coefficient group for a homology theory in S, and let $\rho(t)$ be a polynomial in t with coefficients in $\mathfrak{g}$; ρ is to be thought of as an operator on the chains and cycles of S. Consider an h-dimensional cycle Γ_h such that $\rho(t)\Gamma_h=0$. Such cycles will be called ρ-cycles; in their totality they form an additive group $\mathfrak{H}_h^\rho$. A ρ-cycle Γ_h which is the boundary of a ρ-chain is "ρ-homologous to zero," $\Gamma_h \underset{\rho}{\sim} 0$; denote the totality of these by $\mathfrak{H}_h^0{}^\rho$. The group $H_h^\rho=\mathfrak{H}_h^\rho-\mathfrak{H}_h^0{}^\rho$ is a "special homology group" characteristic of (S, t) and is, in fact, a type invariant of t. Of particular interest are the groups H^σ, $(\sigma=1+t+\cdots+t^{m-1})$, and H^δ, $(\delta=1-t)$, since they can be explicitly determined in a variety of cases and bear certain useful reciprocal relations to each other (see [29]).*

I have recently examined at some length [36] the case in which m is a prime and $S_n=S$ is assumed to be sphere-like in the sense that its dimension is n and that it has the same homology groups as an n-sphere. It can be shown that if H_h^ρ is modified by taking for $\mathfrak{H}_h^0{}^\rho$ the ρ-cycles which are ρ-homologous to zero modulo L, where L is the totality of fixed points, then if any group in the sequence

$$H_n^\rho,\ H_{n-1}^\rho,\ H_{n-2}^\rho,\ \cdots,\ H_0^\rho, \qquad \rho = \sigma \text{ or } \delta,$$

vanishes, all those which follow it also vanish. Let $r=r(\mathfrak{g})$ be the dimension of the first group to vanish; I have shown that if $\mathfrak{g}=\mathfrak{m}$, the integers reduced modulo m, the modulo m dimension† of L is r; and if $r>0$, the Betti numbers (mod m) of L are the same, both locally and in the large, as those of an r-sphere; if $r=0$, L consists of two points. If S_3 is an actual three-sphere, then L is homeomorphic to an

* Strictly speaking, the relation $\rho\Gamma=0$ is not a topological one and in fact has little meaning unless S is a complex and t merely permutes the cells of S among themselves. For situations of a much more general nature, the theory of the σ- and δ-cycles and of the corresponding special homology groups has been worked out in my paper [36]. See also [28] of Richardson, who first introduced the special groups relative to an arbitrary (but fixed) polynomial ρ.

† Alexandroff [2].

actual r-sphere, ($r=2$ or 1), or consists of two points;* if t preserves orientation, $r=1$ unless L is null; if t reverses orientation, $r=0$ or 2, and m can only equal 2.

The methods which lead to these results will apply with slight modifications to the case in which m is composite, although the details are as yet unpublished. I must leave open, however, the problem of determining, in the general case, the dimensional and homological properties of L relative to coefficient groups other than m.† In case S_n is assumed to be a simplicial complex and t is also simplicial (carrying simplexes into simplexes), the results of the preceding paragraph do hold equally well when $\mathfrak{g}$ is say the group of integers. In addition, if t preserves orientation, it can now be shown that the dimension of L is of the same parity as n. Whether or not this is true for the non-simplicial case, even if "dimension" is taken to mean "dimension modulo m," I have as yet not determined.‡

If we no longer assume that S_n is sphere-like, but that it is still fairly regular locally, then although it is no longer true that L must have the homology groups of an r-sphere, locally the situation remains the same as before, and one can show that L possesses to the same degree as S_n whatever regularity can be described by modulo m homology theory. If S_n is an orientable simplicial manifold modulo m, and if t is simplicial and L a subcomplex of S_n, then L consists of a finite number of non-intersecting orientable manifolds (mod m), and the dimension of each is of the same parity as n.§

6. There is, as we shall see, a certain amount of interest in knowing the maximum number of components which L may have under certain conditions. Let us take m to be a prime and assume as above that S_n is an orientable simplicial manifold modulo m and that L is a subcomplex, t simplicial. We may write $L=L_p^1+L_p^2+\cdots+L_p^k+M$, where the L's are non-intersecting manifolds of dimension p and M is the sum of manifolds of dimension not p. I shall establish a certain

* The proof of the statement about S_3 depends on certain topological characterizations of one-dimensional manifolds by Alexandroff [3] and of higher dimensional manifolds by Wilder [39]. Only part of the theorem is stated and proved in [36]; the complete proof will be given elsewhere.

† Partial results are obtained in [36] when $\mathfrak{g}$ is the additive group of rational numbers.

‡ Except for the theorem [36] that if m is prime, the dimension of L can be zero only if n is even.

§ The dimension of L need not be of the same parity as n if S_n is non-orientable. Moreover, when $m=2$, simple examples show that L can be non-orientable when S_n is orientable; this however does not contradict our theorem, since orientability cannot be determined by modulo two topology.

relation between k_p and the group $\mathfrak{H}^{\sigma}_{p+1}$ of §5 (coefficients in $\mathfrak{m}$). Since m is a prime every homology group H is the direct sum of a number τ of cyclic groups of order m; τ will be called the rank of $\mathfrak{H}$. Let τ^{σ}_{p+1} = rank $\mathfrak{H}^{\sigma}_{p+1}$, and let $\tilde{R}_p$ be the maximum number of non-intersecting p-cycles in S linearly independent with respect to homologies. Then the relation in question is

$$k_p \leqq \tau^{\sigma}_{p+1} + \tilde{R}_p. \tag{1}$$

To prove this, choose in each L^i a p-cycle Γ^i which is not homologous to zero in L. It may happen that some or all of the Γ's are independent with respect to homologies in S_n. Suppose the Γ's so named that the first ν of them constitute a maximal set independent in S_n. Let the remaining Γ's now be denoted by $\Delta^1, \cdots, \Delta^{\mu}$, $(\nu+\mu=k_p)$. There is a homology relation between each Δ^i and the cycles $\Gamma^1, \cdots, \Gamma^{\nu}$, say

$$F(A^i_{p+1}) = b^i\Delta^i + \sum_{j=1}^{\nu} a^i_j \Gamma^j, \qquad b^i \neq 0;\ i = 1, \cdots, \mu, \tag{2}$$

where F is the boundary operator. Now δA^i_{p+1} (see §5) is a cycle since $F\delta A = \delta FA = FA - tFA = 0$. Moreover

$$\sigma(\delta A) = (1 + t + \cdots + t^{m-1})(1 - t)A = (1 - t^m)A = 0.$$

Therefore δA may be considered as an element in $\mathfrak{H}^{\sigma}_{p+1}$. Let $C^1, \cdots, C^{\tau}$, $(\tau = \tau^{\sigma}_{p+1})$, be a basis for H^{σ}_{p+1}. Then we may write

$$\delta A^i_{p+1} \underset{\sigma}{\sim} \sum_{j=1}^{\tau} c_j{}^i C^j, \qquad i = 1, \cdots, \mu. \tag{3}$$

I assert that $\mu \leqq \tau$. For suppose $\mu > \tau$. Then there exists a linear relation among the forms in the right of (3) with coefficients $g_1, \cdots, g_{\mu}$ not all zero. Hence from (3) we have

$$\sum_i g_i \delta A^i_{p+1} \underset{\sigma}{\sim} 0.$$

This implies the existence of a chain B_{p+2} such that $\sigma B = 0$ and $FB = \sum g_i \delta A^i_{p+1}$. The relation $\sigma B = 0$ implies the existence* of chains X_{p+2} and X^L_{p+2}, the latter in L, such that $B = \delta X + X^L$. Thus $F(\delta X) + FX^L = \delta(\sum g_i A^i_{p+1})$. Now if $\sum g_i A^i_{p+1}$ contains a cell in L, that cell will appear in $\delta(\sum g_i A^i_{p+1})$ with coefficient zero. Similarly, $F(\delta X) = \delta FX$ contains no cell of L. Hence $FX^L = 0$, and the bounding relation just written takes the form $F(\delta X_{p+2}) = \sum g_i \delta A^i_{p+1}$. Let

* [36], p. 141.

$Z_{p+1} = FX_{p+2} - \sum g_i A^i$. Then $\delta Z = 0$, so that we may write $Z_{p+1} = \sigma U_{p+1} + U^L_{p+1}$, $(U^L \subset L)$.* Now

$$
\begin{aligned}
0 = FFX &= F\Big(\sum g_i A^i + Z\Big) \\
&= F\Big(\sum g_i A^i\Big) + F\sigma U + FU^L \\
&= \sum_i g_i \Big(b^i \Delta^i + \sum_j a_j^i \Gamma^j \Big) + F\sigma U + FU^L .
\end{aligned} \tag{4}
$$

Since the Δ's, the Γ's, and $F(U^L)$ are cycles in L, $F\sigma U$ is also such a cycle. Let E_p be a simplex occurring in $F\sigma U$ with a non-zero coefficient; E_p must occur in the boundary of at least one of the chains $U, tU, \cdots, t^{m-1}U$, say

$$
Ft^j U = yE_p + \cdots , \qquad y \not\equiv 0,
$$

where the dots denote a chain not containing E_p. We then have

$$
\sigma FU = t^j(\sigma FU) = \sigma(Ft^j U) = y\sigma E_p + \cdots .
$$

Since E_p is invariant under t, it follows that $\sigma E_p = mE_p = 0 \pmod m$, and we conclude that $\sigma FU = 0$. Thus (4) becomes

$$
\sum_i g_i \Big(b^i \Delta^i + \sum_j a_j^i \Gamma^j \Big) = - FU^L ;
$$

hence the expression on the left is homologous to zero in L. Therefore the coefficients of the Δ's and Γ's must be zero. Since $b^i \not\equiv 0$ we conclude that $g_i = 0$, $(i = 1, \cdots, \mu)$, which is impossible. This completes the proof of the inequality $\mu \leqq \tau$; and since $\nu \leqq \tilde{R}_p$, (1) is now established.

Consider the matrix a_j of (2), §5, defined now relative to the coefficient group $\mathfrak{m}$. Let q_j^δ be the rank of $e - a_j$, and q_j^σ the rank of $e + a_j + a_j^2 + \cdots + a_j^{m-1}$ (e is the identity matrix). It has been shown elsewhere [29] that

$$
\tau^\sigma_{p+1} \leqq R_{p+1} + R_{p+2} + \cdots + R_n - q^\sigma_{p+1} - q^\delta_{p+2} - \cdots - q^\rho_n , \qquad \rho = \delta \text{ or } \sigma , \tag{5}
$$

where the R's are the modulo m Betti numbers of S_n. If we combine (5) with (1), omitting the q's, we obtain

$$
k_p \leqq \tilde{R}_p + R_{p+1} + \cdots + R_n . \tag{6}
$$

Thus we have the result that *there exists, at least if m is prime, an upper bound for the number of p-dimensional components of L which is*

* [36], p. 141.

independent of the period m and, in fact, depends only on the topological structure of S. Notice that if $m=2$, orientation can play no part in the argument, and (6) then holds equally well for non-orientable manifolds.

Since $\tilde{R}_0=R_0=1$, we have the result that there can be at most $\sum_0^n R_i$ isolated fixed points. Suppose $S=S_2$ is an orientable closed surface of genus p. Then

$$\tilde{R}_1 = p, \quad R_0 = R_2 = 1, \quad R_1 = 2p.$$

Hence there can be at most $p+1$ pointwise invariant curves and $2p+2$ isolated invariant points.

Let $x_0, \cdots, x_s$ be homogeneous coordinates for a complex projective space P_s, and let V be an algebraic variety defined by equating to zero a finite number of polynomials in $x_0, \cdots, x_s$ with real coefficients. Let us assume that, considered as a real locus in a space of $2s$ real dimensions, V is a manifold modulo 2. The transformation $t: (x_0, \cdots, x_s) \to (\bar{x}_0, \cdots, \bar{x}_s)$, where $\bar{x}$ is the complex conjugate of x, is of period two and induces a homeomorphism of V into itself, the fixed points of which constitute the *real* folds or branches of V. Formula (6) then gives an upper bound for the number of connected real branches which V may have. In particular, suppose V is a real algebraic surface. If V has no algebraic singularities, the manifold condition is satisfied (cf. Zariski [41], p. 102), and we can say that there are at most $\tilde{R}_2+R_2+1$ connected real branches (the R's refer of course to V considered as a manifold of four real dimensions). If V is a real algebraic curve, it need not be restricted with regard to algebraic singularities, since V may now be replaced by its Riemann surface which is in any case a manifold. We obtain then the theorem of Harnack that a real algebraic curve of genus p can have at most $p+1$ real branches.*

* We have here taken it for granted that V can be subdivided into simplexes such that the conditions stated at the beginning of §6 are satisfied. There is no difficulty in showing that such is the case by means of the methods used by Brown and Koopman [7] in connection with the triangulation of analytic loci. As a matter of fact the results stated hold equally well for non-simplicial spaces and triangulations, and the assumptions used were only a convenience for exposition. It seems likely also that the condition that V be a manifold could also be dispensed with.

It should be pointed out that little is known about the modulo two Betti numbers of algebraic varieties whereas the non-modular Betti numbers are related to the birational invariants of V. It would therefore be of interest to find upper bounds for k_p in terms of Betti numbers of the latter type rather than the former.

The transformation $x_i \to \bar{x}_i$ was used by Lefschetz [15] in determining the number of real folds of a real abelian variety.

CYCLIC MANIFOLDS

7. We have seen that when S is a simple closed curve, the types of periodic transformations can be explicitly enumerated (§5). Let us consider the problem of enumeration when S is an odd dimensional sphere S_{2n-1}, and t is without fixed points. To make matters as simple as possible, let us suppose that t is a rotation. We may assume that S lies in a euclidean space E_{2n}. If a euclidean coordinate system in E_{2n} is properly chosen, t will be represented by the equations (1) (with the X's identically zero). The θ's are multiples of $2\pi/m$, say

$$\theta_{2i-1} = a_i \cdot 2\pi/m, \qquad 1 \leqq a_i \leqq m-1; \; i = 1, \cdots, n.$$

The a's will be called indices of t; *we shall assume that they are prime to* m. If t' is a second rotation of S with period m, and if t' has a set of indices $b_1, \cdots, b_n$ which, except for order and sign, are the same as $a_1, \cdots, a_n$, then t and t' are topologically equivalent.* Whether or not the converse is true remains, I believe, an open question of considerable interest. If has recently been shown† that the converse is true in a *combinatorial* sense. That is, if for S there exist cellular subdivisions Σ and Σ' identical in structure, with Σ invariant under t, Σ' under t', and if there is a homeomorphism τ of S into itself such that $\tau\Sigma = \Sigma'$ and such that $t' = \tau t \tau^{-1}$, then the a's and b's are the same modulo m except for order and sign. But if one assumes only topological equivalence, then the most that is known at present is that

$$a_1 a_2 \cdots a_n \equiv \pm\, b_1 b_2 \cdots b_n \pmod{m}.$$

This was first proved by de Rham [24] by means of the theory of looping coefficients. One can also establish this relation in a very elementary manner by means of our special homologies. For this purpose three simple lemmas are sufficient. Let us assume for the moment that S possesses a simplicial subdivision which is invariant under t. By a natural sort of extension, t may be thought of as carrying chains into chains in such a way as to preserve boundary relations. Taking as coefficient group either the integers or the integers modulo m, we have the following lemmas:

LEMMA 1. *If* δX *is a* σ*-cycle, then* FX *is a* δ*-cycle; if* $\delta X \underset{\sigma}{\sim} 0$ *then* $FX \underset{\delta}{\sim} 0$. *The same holds with* σ, δ *interchanged.*‡

LEMMA 2. *If* X *is a chain, then* δX *and* $(1-t^a)X$ *are* σ*-chains; if they are cycles, then* $(1-t^a)X \sim a(\delta X)$.

* In fact they are equivalent under an orthogonal transformation.

† de Rham [25]; see also Reidemeister [23], Franz [11].

‡ See [36], p. 143.

PROOF. Since $\sigma=\sigma t=\sigma t^a$, we have $\sigma(\delta X)=\sigma(1-t^a)X=0$, which proves the first assertion. To prove the second, let Y be a chain bounded by δX. Then $\delta FY=\delta\delta X$. Since $\sigma\delta=0$, we may write $\delta\delta X\underset{\sigma}{\sim}0$, or $(1-t)\delta X\sim 0$, or $t\delta X\underset{\sigma}{\sim}\delta X$, and on repeating the argument we have $t^i\delta X\underset{\sigma}{\sim}\delta X$. Hence

$$(1-t^a)X=(1+t+\cdots+t^{a-1})\delta X\underset{\sigma}{\sim}a\delta X.$$

LEMMA 3. *If X_0 is a 0-chain consisting of a single 0-simplex with coefficient ± 1, and if $p\delta X_0\sim 0$, then $p\equiv 0 \pmod m$.*

PROOF. Every σ-chain is of the form δV ([36], p. 141); hence there exists a one-chain Y_1 such that $F\delta Y_1=p\delta X_0$. Let $Z_0=FY_1-pX_0$. Then $\delta Z_0=0$; hence $Z_0=\sigma U_0$ ([36], p. 141). Now the sum of the coefficients of FY_1 is zero (this being the case for the boundary of each individual one-simplex). Moreover the sum of the coefficients of $Z_0=\sigma U_0$ is* zero modulo m, whereas the sum of the coefficients of pX_0 is $\pm p$. Hence $p\equiv 0 \pmod m$.

Suppose now that t is a rotation with indices $a_1, \cdots, a_n$ prime to m. It is easy to show (see [25], p. 741, footnote) that there exist chains $X_0, \cdots, X_{2n-1}$ such that (i) $\sigma X_{2n-1}=\pm S_{2n-1}$; (ii) X_0 is a single vertex with coefficient ± 1; and (iii)

$$\begin{aligned} FX_{2h-1} &= (1-t^{a_h})X_{2h-2}, & h &= 1, \cdots, n, \\ FX_{2h} &= \sigma X_{2h-1}, & h &= 1, \cdots, n, \end{aligned} \tag{1}$$

where $\alpha_h\equiv a_h^{-1} \pmod m$. Let us call any set of chains satisfying these three conditions a normal set with indices a_i. Suppose there is a second normal set Y with indices b_i. Then

$$\prod a_i \equiv \pm \prod b_i \pmod m. \tag{2}$$

For we have $S_{2n-1}=\pm\sigma X_{2n-1}=\pm\sigma Y_{2n-1}$. Hence $\pm\sigma X_{2n-1}\underset{\sigma}{\sim}\sigma Y_{2n-1}$; and from Lemma 1 and (1) we have

$$\pm(1-t^{a_n})X_{2n-2}\underset{\sigma}{\sim}(1-t^{\beta_n})Y_{2n-2}, \qquad \beta_n=b_n^{-1}.$$

Hence from Lemma 2, $\pm(1-t)\alpha_nX_{2n-2}\underset{\sigma}{\sim}(1-t)\beta_nY_{2n-2}$, and from Lemma 1, $\pm\sigma\alpha_nX_{2n-3}\underset{\delta}{\sim}\sigma\beta_nY_{2n-3}$. Again, from Lemma 1 then Lemma 2, we obtain $\pm\alpha_n\alpha_{n-1}\delta X_{2n-4}\sim\beta_n\beta_{n-1}\delta Y_{2n-4}$ and finally

$$\pm\alpha_n\cdots\alpha_0\delta X_0\underset{\sigma}{\sim}\beta_n\cdots\beta_0\delta Y_0.$$

* To each vertex E_0^i there are associated two 0-chains $+E_0^i$ and $-E_0^i$. If $tE_0^i=E_0^j$, it follows, from the condition that t preserve bounding relations, that $t(+E_0^i)=+E_0^j$. Thus $(1+t+\cdots+t^{m-1})(+E_0^i)=+E_0^j+E_0^k+\cdots$; hence the sum of the coefficients of σE_0^i is m.

From Lemma 2 we conclude that $\pm\prod\alpha_i \equiv \prod\beta_i \pmod m$, and from this (2) follows.

Suppose now that t' is a rotation of S_{2n-1} with indices $b_1, \cdots, b_n$. Suppose τ is a homeomorphism of S with itself such that $t=\tau^{-1}t'\tau$. Imagine first that τ carries simplexes into simplexes. If $Y_0, \cdots, Y_{2n-1}$ is a normal set relative to t', with indices b_i, then $\tau^{-1}Y_0, \cdots, \tau^{-1}Y_{2n-1}$ is a normal set relative to t with indices still b_i. Hence $\prod b_i \equiv \pm\prod a_i \pmod m$. In case τ is not simplicial, the images of the Y's under τ are no longer chains of the complex S but will become so after a suitable deformation. The deformation can in fact be so chosen that the new Y's again constitute a normal set for t with indices b_i.

The larger questions connected with the equivalence problem for rotations are of fundamental interest in topology. It has been known for some time that two manifolds, which are identical so far as the homology invariants and the fundamental group are concerned, need not be homeomorphic (J. W. Alexander [1]). Consider a space S and a periodic t of S into itself. The so-called modular space $S^{(t)}$ is constructed by identifying points which are images of each other under powers of t. If S is a sphere of an odd number of dimensions and t a rotation of period m and with indices prime to m, then $S^{(t)}$ is called a cyclic manifold of order m. It can be shown (de Rham [24]) that two cyclic manifolds $S^{(t)}$ and $S^{(t')}$ of order m have the same Betti numbers, the same torsion coefficients, and the same fundamental group. Moreover, a necessary and sufficient condition that $S^{(t)}$ and $S^{(t')}$ be homeomorphic is that there exist an h prime to m such that t^h and t' are topologically equivalent.* Suppose the indices of t and t' are a_i and b_i. Those of t^h are then ha_i. Hence by $S^{(t)}$ and $S^{(t')}$ they can be homeomorphic only if the relation $\pm\prod a_i \equiv h^n b_i \pmod m$ has a solution in h (cf. de Rham [24]). Suppose for example that S is a three-sphere, that $m=5$, and that the indices are (1, 1) for t and (1, 2) for t'. Then $S_3^{(t)}$ and $S_3^{(t')}$ are not homeomorphic, since the equation $\pm 1 \equiv 2h^2 \pmod 5$ has no solutions.†

The problem of finding new topological invariants of manifolds is of course a fundamental one. In focusing attention on cyclic manifolds we have shown the existence of invariants which are independent of the ordinary ones; they are implicit in the relation (2). The existence of a solution for (2) is, however, not sufficient to ensure

* For a given S, the modular space $S^{(t)}$ is a type invariant of t and the special groups $\mathfrak{H}^\sigma, \mathfrak{H}^\delta$ (§5) are closely related to the ordinary homology groups of $S^{(t)}$. These relations have proved to be particularly useful for determining certain relative homology characters for cyclic and symmetric products (see [29]).

† This is essentially the example of Alexander [1].

homeomorphism. If one is content to ask, when are two cyclic manifolds equivalent in a strictly combinatorial sense, one finds a complete answer in recent noteworthy papers of Reidemeister [23], de Rham [25], and Franz [11]. Whether or not these combinatorial results are really topological remains to be determined.

Bibliography

1. J. W. Alexander, *Note on two three-dimensional manifolds with the same group*, Transactions of this Society, vol. 20 (1919), pp. 339–342.

2. P. Alexandroff, *Dimensionstheorie*, Mathematische Annalen, vol. 106 (1932), pp. 161–238.

3. P. Alexandroff, *On local properties of closed sets*, Annals of Mathematics, vol. 36 (1935), pp. 1–35.

4. G. D. Birkhoff and P. A. Smith, *Structure analysis of surface transformations*, Journal de Mathématique, (3), vol. 7 (1928), pp. 345–379.

5. R. Brauer, *Sur les invariants intégraux des variétés des groupes de Lie simples clos*, Comptes Rendus de l'Académie des Sciences, vol. 201 (1935), pp. 419–421.

6. L. E. J. Brouwer, *Die Theorie der endlichen kontinuierlichen Gruppen unabhängig von den Axiomen von Lie*, Mathematische Annalen, vol. 67 (1909), pp. 246–267; and vol. 69 (1910), pp. 181–203.

7. A. B. Brown and B. O. Koopman, *On the covering of analytic loci by complexes*, Transactions of this Society, vol. 34 (1932), pp. 231–251.

8. E. Cartan, *Sur les invariants intégraux de certains espaces homogènes clos et les propriétés topologiques de ces espaces*, Annales de la Société Polonaise de Mathématique, vol. 8 (1929), pp. 181–225.

9. E. Cartan, *La topologie des espaces représentatifs des groupes de Lie*, L'Enseignement Mathématique, vol. 35 (1936), pp. 177–200.

10. E. Cartan, *La Théorie des Groupes Finis et Continus et l'Analyse Situs*, Mémorial des Sciences Mathématiques, vol. 42, 1930.

11. W. Franz, *Über die Torsion einer Überdeckung*, Journal für die reine und angewandte Mathematik, vol. 173 (1935), pp. 245–254.

12. H. Freudenthal, *Einige Sätze über topologische Gruppen*, Annals of Mathematics, (2), vol. 34 (1936), pp. 46–56.

13. W. Hurewicz, *Beiträge zur Topologie der Deformationen*. IV, Proceedings, Akademie van Wetenschappen, Amsterdam, vol. 39 (1936), pp. 215–224.

14. B. von Kerékjártó, *Über die endlichen topologischen Gruppen der Kugelfläche*, Proceedings, Akademie van Wetenschappen, Amsterdam, vol. 22 (1919), pp. 568–569.

15. S. Lefschetz, *On the real folds of abelian varieties*, Proceedings of the National Academy of Sciences, vol. 5 (1919), pp. 102–106.

16. D. Montgomery and L. Zippin, *Periodic one-parameter groups in three-space*, Transactions of this Society, vol. 40 (1936), pp. 24–36.

17. D. Montgomery and L. Zippin, *Translation groups of three-space*, American Journal of Mathematics, vol. 59 (1937), pp. 121–128.

18. J. von Neumann, *Die Einführung analytischer Parameter in topologischen Gruppen*, Annals of Mathematics, (2), vol. 34 (1933), pp. 170–190.

19. J. Nielsen, *Topologischer Beweis eines Satzes von Wiman*, Matematisk Tidsskrift, (B), (1936), pp. 11–24.

20. J. Nielsen, *Die Strukturen periodischer Transformationen von Flächen*, Danske Videnskabernes Selskab, Mathematisk-Fysiske Meddeleleser, vol. 15.

21. L. Pontrjagin, *Sur les groupes topologiques compacts et le cinquième problème de M. Hilbert*, Comptes Rendus de l'Académie des Sciences, vol. 198 (1934), pp. 238–240.

22. L. Pontrjagin, *Sur les nombres de Betti des groupes de Lie*, Comptes Rendus de l'Académie des Sciences, vol. 200 (1935), pp. 1277–1280.

23. K. Reidemeister, *Homotopieringe und Linsenräume*, Abhandlungen aus dem Mathematischen Seminar der Hamburgischen Universität, vol. 11 (1935), pp. 102–109.

24. G. de Rham, *Sur l'analyse situs des variétés à n dimensions*, Journal de Mathématique, (9), vol. 10 (1931), pp. 115–200.

25. G. de Rham, *Sur les nouveaux invariants topologiques de M. Reidemeister*, Recueil Mathématique (Moscow), new series, vol. 1 (1936), pp. 737–742.

26. M. Richardson, *On the homology characters of symmetric products*, Duke Mathematical Journal, vol. 1 (1935), pp. 50–69.

27. M. Richardson, *The relative connectivities of symmetric products*, this Bulletin, vol. 41 (1935), pp. 528–534.

28. M. Richardson, *Special homology groups*, Proceedings of the National Academy of Sciences, vol. 24 (1938), pp. 21–23.

29. M. Richardson and P. A. Smith, *Periodic transformations of complexes*, Annals of Mathematics, (2), vol. 60 (1938).

30. J. F. Ritt, *On certain real solutions of Babbage's functional equation*, Annals of Mathematics, (2), vol. 17 (1915), pp. 113–122.

31. P. A. Smith, *Properties of group manifolds*, Proceedings of the National Academy of Sciences, vol. 17 (1931), pp. 674–675.

32. P. A. Smith, *The topology of involutions*, Proceedings of the National Academy of Sciences, vol. 19 (1933), pp. 612–618.

33. P. A. Smith, *The fundamental group of a group manifold*, Annals of Mathematics, (2), vol. 36 (1935), pp. 210–229.

34. P. A. Smith, *Topological foundations in the theory of continuous transformation groups*, Duke Mathematical Journal, vol. 2 (1936), pp. 246–279.

35. P. A. Smith, *Manifolds with abelian fundamental groups*, Annals of Mathematics, (2), vol. 37 (1936), pp. 526–533.

36. P. A. Smith, *Transformations of finite period*, Annals of Mathematics, (2), vol. 39 (1938), pp. 127–164.

37. F. Steiger, *Die maximalen Ordnungen periodischer topologischer Abbildungen geschlossener Flächen in sich*, Commentarii Mathematici Helvetici, vol. 8 (1935), pp. 48–69.

38. R. J. Walker, *The Betti numbers of cyclic products*, this Bulletin, vol. 42 (1936), pp. 709–714.

39. R. L. Wilder, *Generalized closed manifolds in n-space*, Annals of Mathematics, (2), vol. 35 (1934), pp. 876–903.

40. A. Wiman, *Über die hyperelliptischen Curven und diejenigen vom Geschlecht $p=3$ welche eindeutige Transformationen in sich zulassen*, Svenska Vetenskapsakademiens Hanglingar, vol. 2 (1895).

41. O. Zariski, *Algebraic Surfaces*, Ergebnisse der Mathematik und ihrer Grenzgebiete, (3), vol. 5, 1935.

BARNARD COLLEGE,
COLUMBIA UNIVERSITY

NOTE ON AN ELEMENTARY PROBLEM OF INTERPOLATION

P. ERDÖS AND G. GRÜNWALD

The unique polynomial of degree $n-1$ assuming the values $y_1, y_2, \cdots, y_n$ at the abscissas $x_1, x_2, \cdots, x_n$, respectively, is given by the Lagrange interpolation formula

$$L_n(x) = y_1 l_1(x) + y_2 l_2(x) + \cdots + y_n l_n(x), \tag{1}$$

where

$$l_k(x) = \frac{\omega(x)}{\omega'(x_k)(x - x_k)}, \qquad k = 1, 2, \cdots, n, \tag{2}$$

(fundamental polynomials of the Lagrange interpolation) and the polynomial $\omega(x)$ is defined by

$$\omega(x) = c(x - x_1)(x - x_2) \cdots (x - x_n), \tag{3}$$

where c denotes an arbitrary constant not equal to zero.

In this note we prove the following theorem:

THEOREM. *In the Lagrange interpolation formula let* $x_k = x_k^{(n)} = \cos(2k-1)\pi/2n = \cos\theta_k^{(n)}$, $(k = 1, 2, \cdots, n)$, *which implies* $\omega(x) = T_n(x) = \cos(n \arccos x) = \cos n\theta$ (*Tschebycheff polynomial*). *Then*

$$|l_k^{(n)}(x)| = \left|\frac{\omega(x)}{\omega'(x_k)(x - x_k)}\right| < \frac{4}{\pi}, \qquad -1 \leqq x \leqq +1, \tag{4}$$

for all n *and* k, *and furthermore*

$$\lim_{n\to\infty} |l_1^{(n)}(+1)| = \lim_{n\to\infty} |l_n^{(n)}(-1)| = \frac{4}{\pi}. \tag{5}$$

In this connection Fejér* proved for all n, k, and x, $(-1 \leqq x \leqq +1)$,

$$|l_k^{(n)}(x)| < 2^{1/2}. \tag{6}$$

Of course (5) implies that inequality (4) is the best possible in the following sense: For any $\epsilon > 0$ there exist values of n, k, and x,

* L. Fejér, *Lagrangesche Interpolation und die zugehörigen konjugierten Punkte*, Mathematische Annalen, vol. 106 (1932), pp. 1–55; see pp. 10, 11. This paper will hereafter be referred to as L.

$(-1 \leqq x \leqq +1)$, such that

$$|l_k^{(n)}(x)| > \frac{4}{\pi} - \epsilon. \tag{7}$$

Our proof depends upon the Hermite interpolation formula which gives the unique polynomial $H(x)$ of degree $2n-1$ satisfying the conditions

$$H(x_k) = y_k, \qquad H'(x_k) = y_k', \qquad k = 1, 2, \cdots, n, \tag{8}$$

where the y_k and y_k' are given numbers. It is easy to show that*

$$H(x) = \sum_{k=1}^{n} y_k v_k(x)\{l_k(x)\}^2 + \sum_{k=1}^{n} y_k'(x - x_k)\{l_k(x)\}^2, \tag{9}$$

where

$$v_k(x) = 1 - (x - x_k)\frac{\omega''(x_k)}{\omega'(x_k)}, \tag{10}$$

$$\sum_{k=1}^{n} v_k(x)\{l_k(x)\}^2 \equiv 1. \tag{11}$$

For the Tschebycheff abscissas we have

$$v_k(x) = v_k^{(n)}(x) = \frac{1 - xx_k^{(n)}}{1 - (x_k^{(n)})^2}, \qquad x_k^{(n)} = \cos(2k-1)\pi/2n. \tag{12}$$

Fejér proved (6) by aid of the simple inequality $v_k^{(n)}(x) \geqq 1/2$.†

We also need the following result due to M. Riesz:‡

LEMMA. *A trigonometric polynomial of degree $n-1$ assumes the maximum of its absolute value at a point whose distance from any of the roots of this trigonometric polynomial is not less than $\pi/[2(n-1)]$.*

We are now in position to prove the theorem. For $n=1$ and $n=2$

$$|l_1^{(1)}(x)| = 1, \quad |l_1^{(2)}(x)| = \left|\frac{(\sin \pi/4)\cos 2\theta}{2(\cos\theta - \cos\pi/4)}\right|$$

$$= \frac{1}{2^{1/2}}\left|\cos\theta + \frac{1}{2^{1/2}}\right| < \frac{1 + 2^{1/2}}{2} < \frac{4}{\pi},$$

* L. Fejér, *Weierstrasssche Approximation, besonders durch Hermitische Interpolation*, Mathematische Annalen, vol. 102 (1930), pp. 707–725.

† See L, p. 5.

‡ M. Riesz, *Eine trigonometrische Interpolationsformel und einige Ungleichungen für Polynome*, Jahresbericht der Deutschen Mathematiker-Vereinigung, vol. 23 (1914), pp. 354–368; pp. 363–364.

$$|l_2^{(2)}(x)| < \frac{1}{2^{1/2}}\left|\cos\theta - \frac{1}{2^{1/2}}\right| < \frac{1+2^{1/2}}{2} < \frac{4}{\pi}.$$

Thus we have to consider only the case $n \geqq 3$. For the Tschebycheff abscissas we have*

$$(13)\quad l_k^{(n)}(x) = (-1)^{k+1}n^{-1}\sin\theta_k^{(n)}\frac{\cos n\theta}{\cos\theta - \cos\theta_k^{(n)}}, \qquad x = \cos\theta.$$

From (13) it follows that $l_k^{(n)}(\cos\theta)$ is a trigonometric polynomial of degree $n-1$. For $2 \leqq k \leqq n-1$, the roots of $l_k^{(n)}(\cos\theta)$ in $(0, \pi)$ are

$$(14)\qquad \theta_\nu^{(n)} = (2\nu - 1)\frac{\pi}{2n}, \qquad 1 \leqq \nu \leqq n, \nu \neq k,$$

and since $\theta_{\nu+1}^{(n)} - \theta_\nu^{(n)} = \pi/n$, $\theta_1^{(n)} - 0 = \pi - \theta_n^{(n)} = \pi/2n$, $|l_k^{(n)}(\cos\theta)|$ assumes its maximum between $\theta_{k-1}^{(n)}$ and $\theta_{k+1}^{(n)}$. Further it is clear that $|l_1^{(n)}(\cos\theta)|$ and $|l_n^{(n)}(\cos\theta)|$ assume their maxima at $\theta = 0$ and $\theta = \pi$, respectively. Let us consider first $l_1^{(n)}(x)$ and $l_n^{(n)}(x)$. According to the last remark it will be sufficient to find bounds for $|l_1^{(n)}(+1)|$ and $|l_n^{(n)}(-1)|$. From (13) we have

$$|l_1^{(n)}(+1)| = |l_n^{(n)}(-1)| = \frac{\sin\theta_1^{(n)}}{n(1-\cos\theta_1^{(n)})} = \frac{1}{n}\cot\frac{\pi}{4n},$$

whence

$$(15)\qquad \lim_{n\to\infty}|l_1^{(n)}(+1)| = \lim_{n\to\infty}|l_n^{(n)}(-1)| = \frac{4}{\pi}.$$

By differentiation we easily see that $x\cot x$ decreases if x increases so that

$$(16)\quad |l_1^{(n)}(+1)| \leqq |l_1^{(n+1)}(+1)|, \quad |l_n^{(n)}(-1)| \leqq |l_{n+1}^{(n+1)}(-1)|.$$

From (15) and (16) we obtain (7), that is, the second part of the statement.

We now prove that

$$(17)\qquad \max_{-1\leqq x\leqq +1}|l_k^{(n)}(x)| < |l_1^{(n)}(+1)|, \qquad 2 \leqq k \leqq n-1.$$

By (16) it suffices to show that

$$(18)\qquad \max_{-1\leqq x\leqq +1}|l_k^{(n)}(x)| = |l_k^{(n)}(\mu_k)| < |l_1^{(2)}(+1)| = \tfrac{1}{2}(1+2^{1/2}).$$

* See L, p. 5.

In order to prove (18) we show that

$$v_k^{(n)}(\mu_k) > \frac{13}{18}, \qquad 2 \leqq k \leqq n-1. \tag{19}$$

Then (11) furnishes, since $v_k(x) \geqq 0$,

$$1 = \sum_{k=1}^{n} v_k^{(n)}(\mu_k)\{l_k^{(n)}(\mu_k)\}^2 > \frac{13}{18}\{l_k^{(n)}(\mu_k)\}^2,$$

that is,

$$|l_k^{(n)}(\mu_k)| < \left[\frac{18}{13}\right]^{1/2} < (1.4)^{1/2} < 1.2 = \tfrac{1}{2}(1+1.4) < \tfrac{1}{2}(1+2^{1/2}).$$

Let $\mu_k = \cos\phi$, $(0 < \phi < \pi)$. According to the lemma we have $|\phi - \theta_k^{(n)}| < \pi/2n$. On account of (12) it is sufficient to prove

$$\frac{1-\cos\theta_k^{(n)}\cos(\theta_k^{(n)}+\delta)}{1-(\cos\theta_k^{(n)})^2} \geqq \frac{13}{18}, \qquad \delta = \pm\frac{\pi}{2n}. \tag{20}$$

We can assume that $\theta_k^{(n)} < \pi/2$ and $\delta = -\pi/2n$. If we write $\theta_k^{(n)} = 3\mu$, we have $\theta_k^{(n)} + \delta \geqq 2\mu$ and $\cos\mu = t$, $(\frac{1}{2} \leqq t \leqq +1)$, hence

$$\frac{\cos(\theta_k^{(n)}+\delta) - \cos\theta_k^{(n)}}{1-\cos\theta_k^{(n)}} \leqq \frac{\cos 2\mu - \cos 3\mu}{1-\cos 3\mu} = \frac{4t^2+2t-1}{4t^2+4t+1}.$$

The last fraction is $\leqq 5/9$ so that

$$\frac{\cos\theta_k^{(n)}}{1+\cos\theta_k^{(n)}} \cdot \frac{\cos(\theta_k^{(n)}+\delta) - \cos\theta_k^{(n)}}{1-\cos\theta_k^{(n)}} \leqq \frac{5}{18},$$

which is equivalent to (20). This completes the proof.

VICTORIA UNIVERSITY AND
BUDAPEST, HUNGARY

ON THE PEANO CURVE OF LEBESGUE*

I. J. SCHOENBERG

Let $f(t)$ be the *even* continuous function of period two which is defined in the interval $(0, 1)$ as follows: $f(t)=0$ in $(0, 1/3)$, $f(t)=1$ in $(2/3, 1)$, and $f(t)$ is linear in $(1/3, 2/3)$. Our curve is defined by the parametric equations

$$(1)\qquad \begin{aligned} x(t) &= \frac{1}{2}f(t) + \frac{1}{2^2}f(3^2t) + \frac{1}{2^3}f(3^4t) + \cdots, \\ y(t) &= \frac{1}{2}f(3t) + \frac{1}{2^2}f(3^3t) + \frac{1}{2^3}f(3^5t) + \cdots, \quad 0 \leqq t \leqq 1. \end{aligned}$$

The inequalities $0 \leqq f(t) \leqq 1$ imply $0 \leqq x(t) \leqq 1$, $0 \leqq y(t) \leqq 1$, as well as the uniform convergence of both series (1), and hence imply the continuity of $x(t)$, $y(t)$. All there remains to show is that our curve will pass through an arbitrarily given point

$$(2)\quad x_0 = \frac{a_0}{2} + \frac{a_2}{2^2} + \frac{a_4}{2^3} + \cdots, \quad y_0 = \frac{a_1}{2} + \frac{a_3}{2^2} + \frac{a_5}{2^3} + \cdots, \quad a_\nu = 0, 1,$$

of the square $0 \leqq x, y \leqq 1$, whose coordinates are given by their binary expansions. Indeed, let

$$(3)\qquad t_0 = \frac{2a_0}{3} + \frac{2a_1}{3^2} + \frac{2a_2}{3^3} + \cdots + \frac{2a_{k-1}}{3^k} + \frac{2a_k}{3^{k+1}} + \cdots.$$

If $a_0=0$, we have $0 \leqq t_0 \leqq 2/3^2+2/3^3+\cdots = 1/3$, hence $f(t_0)=0$; if $a_0=1$, we have $2/3 \leqq t_0 \leqq 2/3+1/3=1$, hence $f(t_0)=1$. In either case $f(t_0)=a_0$. Similarly $3^k t_0 = \text{even integer} + 2a_k/3 + 2a_{k+1}/3^2 + \cdots$ shows that

$$(4)\qquad f(3^k t_0) = a_k, \qquad k = 0, 1, 2, \cdots.$$

Now (1), (2), and (4) imply $x(t_0)=x_0$, $y(t_0)=y_0$.†

COLBY COLLEGE

* Presented to the Society, April 16, 1938.

† Lebesgue, *Leçons sur l'Intégration*, Paris, 1928, pp. 44–45, defines the functions $x(t)$, $y(t)$, first on Cantor's ternary set T, of points t_0 of the form (3), by means of the equations (2). Having proved their continuity on T, Lebesgue extends their definition throughout $(0, 1)$ by linear interpolation over each one of the denumerable set of open intervals of which the set complementary to T is built up. It should be remarked that our curve (1) coincides with Lebesgue's curve within Cantor's set T but not on the complement of T.

NOTE ON THE CURVATURE OF ORTHOGONAL TRAJECTORIES OF LEVEL CURVES OF GREEN'S FUNCTIONS

J. L. WALSH

The writer has recently established the following theorem:*

THEOREM I. *Let R be a simply connected region of the extended (x, y)-plane whose boundary B contains at least two points. Let $G(x, y)$ be Green's function for R with pole in the point O. Let $\{T\}$ denote the set of orthogonal trajectories to the level curves $G(x, y) = \log r$, $(0 < r < 1)$. The totality of circles each osculating a curve T at O is precisely the set of circles through O and through another fixed point D depending on O and R. There exists no circle that separates both O and D from B.*

The curves T are of course the images of the radii under the conformal mapping of a circle onto R, so that the center of the circle corresponds to O.

In Theorem I, the term "circle" is used in the extended sense, to include straight lines.

In the situation of Theorem I, we shall call D the *conjugate* of O with respect to R. This terminology seems justified, because in the case that B is a line, the point D is the reflection of O in B; and in the case that B is a proper circle, the point D is the inverse of O in B.

It is the object of the present note to establish the following theorem:

THEOREM II. *Under the hypotheses of Theorem* I *the point D may be chosen arbitrarily exterior to R; if D is so chosen there exists at least one point O interior to R whose conjugate with respect to R is D.*

In the proof of Theorem II it is sufficient, by the use of a suitably chosen linear transformation, to treat the special case that D is chosen at infinity.†

THEOREM III. *Let R be an arbitrary limited simply connected region R. There exists interior to R at least one point O whose conjugate D lies at infinity. That is to say, there exists at least one point O of R such that the orthogonal trajectories to the level curves of Green's function with pole at O have at O zero curvature.*

* Proceedings of the National Academy of Sciences, vol. 23 (1937), pp. 166–169.

† Choice of O at infinity yields the theorem: *Let R be a simply connected region, and let D be a point exterior to R. There exists a linear transformation of the plane which carries R into an infinite region for which the image of D is the conformal center of gravity.*

The proof of Theorem III is to be set forth with the aid of several lemmas. If R is a limited simply connected region of the z-plane, the *inner radius* of R with respect to an interior point $O: z=a$ is defined as the unique positive number $r(a)$ such that R can be mapped conformally onto the region $|w|<r(a)$ by a function of the form

$$z - a = w + c_2w^2 + c_3w^3 + \cdots, \qquad |w| < r(a). \tag{1}$$

LEMMA I. *If R is a limited simply connected region of the z-plane, if $O: z=a$ lies interior to R, and if $\delta(a)$ denotes the distance from O to the boundary of R, then we have*

$$4\delta(a) \geqq r(a). \tag{2}$$

Let the function (1) map R onto $|w|<r(a)$. Then the function

$$\zeta = \frac{1}{r(a)}\left[r(a)\omega + c_2[r(a)]^2\omega^2 + c_3[r(a)]^3\omega^3 + \cdots\right], \qquad |\omega| < 1,$$

maps the region R_1 of the ζ-plane, obtained from R by the transformation $\zeta=(z-a)/r(a)$, onto the region $|\omega|<1$. The distance from the point $\zeta=0$ to the boundary of R_1 is $\delta(a)/r(a)$, so inequality (2) follows at once from the Verzerrungssatz.* An immediate conclusion is the following lemma:

LEMMA II. *Under the hypothesis of Lemma* I *the function $r(a)$ approaches zero whenever the point $O: z=a$ approaches the boundary of R.*†

---

* See for instance Pólya-Szegö, *Aufgaben und Lehrsätze aus der Analysis*, vol. 2, Berlin, 1925, p. 27, exercise 152. Lemma II is given there explicitly (p. 19, exercise 112) for the case that R is bounded by an analytic curve.

† Let the function $z=f(w)$ map conformally $|w|<1$ onto R, with $a=f(b)$; then the function $z=\phi(w)=f[(w+b)/(1+\bar{b}w)]$ maps $|w|<1$ onto R with $a=\phi(0)$, and we have $r(a)=|\phi'(0)|=|f'(b)|(1-|b|^2)$, so Lemma I is precisely the inequality

$$4\delta(a) \geqq |f'(b)|(1-|b|^2);$$

Lemma II asserts precisely the conclusion

$$\lim |f'(b)|(1-|b|^2) = 0, \tag{A}$$

as a approaches the boundary of R (or, what is the same thing, as $|b|\to 1$).

Still another method of proof of Lemma II is the following. Instead of keeping the region R fixed and allowing a sequence of interior points $z=a$ to approach the boundary of R, we may study $r(a)$ by keeping the point $z=a$ fixed and allowing the region R to vary by a sequence of translations in such a way that the boundary of the variable R approaches $z=a$. Again by the use of $\phi(w)$, Carathéodory's theory of the conformal mapping of variable regions then yields Lemma II.

These two distinct methods in connection with the study of $\phi(w)$, namely (i) use of various inequalities such as (2), and (ii) Carathéodory's theory of variable regions, can be employed not merely for the proof of (A) but can also be used to study higher derivatives of $f(w)$, as will be shown on another occasion.

The following lemma will also be useful:

LEMMA III. *If $r(a)$ has a relative maximum at the point O: $|z|=a$, then for the corresponding mapping function* (1) *we have $c_2=0$, and the conjugate of O lies at infinity.*

Under the present hypothesis the relation $a_2=0$ is not difficult to establish,* where the inverse of (1) is given by

$$w = (z-a) + a_2(z-a)^2 + a_3(z-a)^3 + \cdots .$$

By the usual formulas for the inversion of the power series (1), we then have $c_2=0$, from which it follows† that the conjugate of O lies at infinity.

We are now in a position to establish Theorem III. The function $r(a)$ is positive and continuous‡ at each point $O: z=a$ of R. Whenever the point O in R approaches the boundary of R, the function $r(a)$ approaches zero, by Lemma II. The function $r(a)$, when suitably defined on the boundary, is then continuous in the corresponding closed region and possesses an absolute maximum in that region. That maximum occurs interior to R and is necessarily a relative maximum of $r(a)$. Theorem III follows from Lemma III.

We now give an example to show that *when D is given the point O of which D is the conjugate need not be unique.* For convenience we choose the situation of Theorem III. Let R be the region of the z-plane formed by the interiors of the two circles

$$\left| z \pm (.99)^{1/2} \right| = 1;$$

the two points $z=\pm i/10$ lie on the boundary of R. In the notation already introduced, we have $r(0)\leqq .4$, by Lemma I. But R contains in its interior the region $\left|z-(.99)^{1/2}\right|<1$, so we have§ $r(.99)^{1/2}\geqq 1$. Thus the point $z=0$ does not furnish an absolute maximum for the function $r(a)$; such an absolute maximum exists interior to R; such an absolute maximum cannot be unique, for if $z=\alpha$ furnishes an absolute maximum, so also does $z=-\alpha\neq\alpha$; and every such absolute maximum has its conjugate at infinity.

In the example just given, the point $z=0$ has its conjugate at infinity even though it does not give an absolute maximum to the function $r(a)$.

---

* Pólya-Szegö, op. cit., p. 19, exercise 113.

† Walsh, loc. cit.

‡ The continuity follows for instance from the formula for $r(a)$ in terms of a fixed $r(b)$; Pólya-Szegö, op. cit., p. 19, exercise 110.

§ Pólya-Szegö, op. cit., p. 21, exercise 121.

A slight modification of the example given shows that if N is chosen arbitrarily, there exists a limited region* having at least N distinct points O whose conjugates D lie at infinity.†

Theorem III becomes false if in the hypothesis the region R is not assumed limited, for the reader may verify that no point O of the region R has its conjugate at infinity if R is the entire plane slit along the positive half of the axis of reals from the point $z=0$ to infinity.

HARVARD UNIVERSITY

A NOTE ON LINEAR FUNCTIONALS

R. P. BOAS, JR.,‡ AND J. W. TUKEY

1. **Introduction.** The knowledge of the general form of linear functionals in a given abstract space§ is of value in many problems. In some cases (notably in the theory of moment problems) the applications are not to the given space, but to its conjugate space; for example, since the general linear functional on L, the space of functions $x=x(t)$ integrable on (0, 1), has the form

$$(1.1)\quad f(x) = \int_0^1 x(t)a(t)dt, \quad a \text{ measurable and essentially bounded,}\|$$

one can solve the moment problem

$$(1.2)\qquad \mu_n = \int_0^1 t^n a(t)dt, \qquad n = 0, 1, 2, \cdots,$$

for essentially bounded functions a.¶ From the point of view of the theory of moment problems, it seems quite fortuitous that there

---

* For the *unlimited* region R: $|y| \leqq b > 0$ of the (x, y)-plane, every point $(x, 0)$ has as conjugate the point at infinity.

† The referee points out that for any region the set of points O whose conjugates D lie at infinity is identical with the set of critical points of the function $r(a)$.

‡ National Research Fellow.

§ We use the terminology of S. Banach, *Théorie des Opérations Linéaires*, Warsaw, 1932.

‖ S. Banach, op. cit., p. 65. The function a is said to be essentially bounded if there is a number M such that $|a(t)| \leqq M$ for almost all t on (0, 1); we denote by $\sup^e |a(t)|$ the greatest lower bound of such numbers M.

¶ S. Banach, op. cit., p. 75.

should exist a space, containing the functions t^n, $(n=0, 1, 2, \cdots)$, on which (1.1) is the general linear functional; and if one attempts to discuss the moment problem (1.2) when a is, for example, continuous or integrable,* one meets the difficulty that

$$(1.3) \qquad f(x) = \int_0^1 x(t)a(t)dt, \quad a \text{ continuous or integrable,}$$

is not the general linear functional on any space for which the general linear functional has yet been determined.

One is thus led to seek the spaces on which a given expression (in particular, (1.3)) can represent the general linear functional. The purpose of this note is to show that (1.3) can represent the general linear functional on no "interesting" space; it follows, by a result of J. W. Tukey,† that necessary and sufficient conditions for the solubility of the moment problem (1.2), with a continuous or integrable, cannot take the form which F. Riesz, E. Helly, and S. Banach have given for other moment problems.‡

It has been shown by I. Gelfand§ that the spaces C and L (whose elements are continuous and integrable functions, respectively, with the classical metrics) are not the conjugate spaces of any Banach spaces. Gelfand's results clearly neither include ours nor are included in them. Our problem is of quite a different character. In Gelfand's problem, a particular Banach space is given, and one asks whether or not it is the conjugate of some Banach space. In the problem of which we consider certain special cases, we are given two vector spaces J and K, without topology, and a functional (j, k) defined on the Cartesian product of J and K, distributive‖ with respect to each argument separately. We ask whether or not a topology (satisfying certain conditions) can be introduced into some linear subset H of J (also satisfying certain conditions) in such a way that H becomes a Banach space with the properties that every linear functional on H (with this topology) has the form $f_k=f_k(h)=(h, k)$, $h \in H$, $k \in K$, and every functional of this form is linear on H. If such a set H exists, a topology is induced in K, possibly with identification of some of the elements of K, by considering K as the conjugate space of H. The topologies in H and K obtained in this way may be entirely different

* In the sense of Lebesgue.

† To be published later.

‡ S. Banach, op. cit., pp. 57, 74–75.

§ I. Gelfand, *Zur Theorie abstrakter Funktionen*, Comptes Rendus (Doklady) de l'Académie des Sciences de l'URSS, vol. 17 (1937), pp. 243–245.

‖ An operation is called distributive if it is additive and homogeneous.

from those which one would naturally expect in those spaces. We note that in the theory of moment problems the topology of the vector spaces involved is quite extraneous, so that our problem arises naturally in connection with that theory.

2. **The theorems.** We first make precise the notion of "general linear functional."

DEFINITION. *Let R be a topological vector space of elements x. Let P be a space of elements p. Let $f_p(x)$ be a functional with domain R, defined for each p in P. We say, "a general linear functional in R is $f_p(x)$," if*

(i) *$f_p(x)$ is a linear functional for every p in P, and*

(ii) *every linear functional $g(x)$ with domain R is identically equal to some $f_p(x)$.*

We do not consider the question of whether or not this representation of the general linear functional in R is redundant; that is, whether $f_p(x) = f_q(x)$ for every x in R is possible with $p \neq q$.

We consider first the functional (1.3) with integrable a.

THEOREM 1. *Let R be a normed vector space whose elements are real-valued functions $x = x(t)$, $(0 \leqq t \leqq 1)$, (with the usual definitions of operations on elements). Let R contain an infinite subset of the functions t^n, $(n = 1, 2, \cdots)$. Then*

$$f_a(x) = \int_0^1 x(t)a(t)dt, \quad a \text{ integrable}, \tag{2.1}$$

is not a general linear functional on R.

We know no necessary and sufficient conditions for (2.1) to be a general linear functional on a space R. It is clear that (2.1) is a general linear functional on any finite-dimensional subspace of B (the space whose elements are essentially bounded measurable functions on (0, 1), with the essential upper bound as norm); another result of the same character is the following theorem:

THEOREM 2. *Let E_n, $(n = 1, 2, \cdots)$, be mutually disjoint measurable subsets of (0, 1), of positive measure. Let $x_n(t)$ be the characteristic function of E_n, $(n = 1, 2, \cdots)$. Let R be the normed space whose elements are all finite linear combinations of the $x_n(t)$, with $\|x\| = \sup_{0 \leqq t \leqq 1} |x(t)|$. Then a general linear functional on R is (2.1).*

It is clear that the representation (2.1) is redundant (in the sense explained above).

For the functional (1.3) with continuous a, our result is the following theorem:

THEOREM 3. *Let R be a vector space of type F,* whose elements are real-valued functions $x=x(t)$, $(0\leqq t\leqq 1)$, (with the usual definitions of operations on elements). Let R contain all of the functions t^n, $(n=0, 1, 2, \cdots)$, and let R have the properties that*

(a) $|x(t)|$ *is in R if $x(t)$ is in R,*

(b) $\lim_{n\to\infty}x_n(t)=\Theta$ *implies* $\lim_{n\to\infty}|x_n(t)|=\Theta$, *where Θ denotes the zero element of R.*

Then

$$f_a(x)=\int_0^1 x(t)a(t)dt, \quad a \text{ continuous}, \tag{2.2}$$

is not a general linear functional on R.

3. **Proof of Theorem** 1. Let R satisfy the conditions of Theorem 1. Suppose, if possible, that (2.1) is a general linear functional on R. Let $y=y(t)\equiv 1$, $(0\leqq t\leqq 1)$. Then, for any x in R,

$$f_y(t)=\int_0^1 x(t)y(t)dt=\int_0^1 x(t)dt$$

exists; that is, x is integrable. Furthermore, x is essentially bounded, since otherwise, by a theorem of Lebesgue,† $\int_0^1 x(t)a(t)dt$ could not exist for every integrable $a(t)$.

If $\lim_{n\to\infty}\|x_n\|=0$, then

$$f_a(x_n)=\int_0^1 x_n(t)a(t)dt\to 0, \qquad n\to\infty,$$

for every integrable a, since f_a is a linear functional. Consequently, by another theorem of Lebesgue,‡

$$\lim_{n\to\infty}\sup_{0\leqq t\leqq 1}{}^0\ |x_n(t)|=0;$$

that is, $\lim_{n\to\infty}\|x_n\|_B=0$, where $\|x\|_B$ is the norm of x in the space B (mentioned above). It follows that any linear functional on B is a linear functional on R.

Let $g=g(x)$ be the functional, constructed by T. H. Hildebrandt,§

* S. Banach, op. cit., chap. 3. We do not use the full force of this hypothesis.

† E. W. Hobson, *The Theory of Functions of a Real Variable and the Theory of Fourier's Series*, vol. 1, 1927, p. 581.

‡ E. W. Hobson, op. cit., vol. 2, 1926, p. 438.

§ T. H. Hildebrandt, *On bounded linear functional operations*, Transactions of this Society, vol. 36 (1934), pp. 868–875; p. 875.

such that g is linear on B (and consequently on R), and

$$g[x(t)] = x(1)$$

for every continuous x. If g had the form (2.1), we should have, in particular,

$$g(t^n) = \int_0^1 t^n a(t)dt = 1 \tag{3.1}$$

for the infinite sequence of values of n for which t^n is in R. But $t^n \to 0$, $(n \to \infty)$, boundedly for $0 \leqq t < 1$, so that $\lim_{n\to\infty} \int_0^1 t^n a(t)dt = 0$, which contradicts (3.1). Consequently (2.1) is not a general linear functional on R.

4. **Proof of Theorem** 2. Let R be the space defined in the hypothesis of Theorem 2. Let f be an arbitrary linear functional on R, and let $c_n = f(x_n)$, $(n = 1, 2, \cdots)$. Since R is a normed vector space, there exists a constant M such that for any real numbers h_i, $(i = 1, 2, \cdots)$, and any integer n,

$$\left| \sum_{i=1}^{n} h_i c_i \right| \leqq M \left\| \sum_{i=1}^{n} h_i x_i \right\| = M \max_{1 \leqq i \leqq n} |h_i|.^* \tag{4.1}$$

Take $h_i = \operatorname{sgn} c_i$, $(i = 1, 2, \cdots)$. Then (4.1) implies that

$$\sum_{i=1}^{n} |c_i| \leqq M, \qquad n = 1, 2, \cdots,$$

and consequently that $\sum_{i=1}^{\infty} |c_i| \leqq M$. Define $a = a(t)$ by

$$a(t) = \begin{cases} c_i \cdot |E_i|^{-1}, & t \in E_i, \\ 0, & t \notin \sum_{i=1}^{\infty} E_i, \end{cases}$$

where $|E|$ denotes the Lebesgue measure of E. Then a is integrable, since

$$\int_0^1 |a(t)|\, dt = \sum_{i=1}^{\infty} \int_{E_i} |c_i| \cdot |E_i|^{-1} dt = \sum_{i=1}^{\infty} |c_i| \leqq M.$$

We have

$$\int_0^1 x_n(t) a(t) dt = c_n = f(x_n), \qquad n = 1, 2, \cdots. \tag{4.2}$$

* S. Banach, op. cit., p. 57.

Since the E_n are mutually disjoint and $f(x)$ is linear, (4.2) implies that

$$f(x) = \int_0^1 x(t)a(t)dt, \qquad x \in R.$$

5. **Proof of Theorem** 3. Let R have the properties specified in Theorem 3, and suppose, if possible, that a general linear functional in R has the form (2.2). Then, as in §3, any x in R is integrable. Define a function $b = b(t)$ by

$$b(t) = \begin{cases} 1, & 0 \leqq t \leqq \frac{1}{2}, \\ 0, & \frac{1}{2} < t \leqq 1. \end{cases}$$

Then

$$g = g(x) = \int_0^1 x(t)b(t)dt$$

is a distributive functional on R but is not linear. For, if g were continuous, we should have, for all x in R and some continuous $b_1(t)$,

$$\int_0^1 x(t)[b(t) - b_1(t)]dt = 0,$$

and in particular

$$\int_0^1 t^n[b(t) - b_1(t)]dt = 0, \qquad n = 0, 1, 2, \cdots; \tag{5.1}$$

but it is well known that (5.1) implies that $b(t) = b_1(t)$ for almost all t.

Since g is not linear, there must exist elements x_n of R, $(n = 1, 2, \cdots)$, such that $\lim_{n\to\infty} x_n = \Theta$, while $g(x_n) \geqq 1$. Let $c(t) = 1$, $(0 \leqq t \leqq 1)$. Since $c(t) \geqq b(t) \geqq 0$, $(0 \leqq t \leqq 1)$, we have, for $n = 1, 2, \cdots$,

$$1 \leqq g(x_n) = \int_0^1 x_n(t)b(t)dt \leqq \int_0^1 |x_n(t)|\, b(t)dt \leqq \int_0^1 |x_n(t)|\, c(t)dt.$$

However, $h(x) = \int_0^1 x(t)c(t)dt$ is a linear functional; hence, by properties (a) and (b),

$$\lim_{n\to\infty} \int_0^1 |x_n(t)|\, c(t)dt = 0.$$

This contradiction establishes Theorem 3.

PRINCETON UNIVERSITY

ON CERTAIN GROUPS OF BIRATIONAL CONTACT TRANSFORMATIONS*

J. M. FELD

1. **Introduction.** Let x_i and u_i, $(i=1, 2, 3)$, be the projective coordinates in plane xu of a point and line, respectively, and y_i and v_i the coordinates of a point and line in plane yv. The planes may be superimposed. Let s be a transformation in mixed coordinates

$$\rho y_i = \phi_i(x, u), \qquad \sigma v_i = \psi_i(x, u), \qquad i = 1, 2, 3, \tag{1}$$

where the ϕ_i and the ψ_i are polynomials homogeneous in the x_i of degrees a and c, respectively, and are also polynomials in the u_i of degrees b and d, respectively. This transformation is said to be birational if from (1) it is possible to obtain the inverse s^{-1} given by

$$\rho' x_i = \phi_i'(y, v), \qquad \sigma' u_i = \psi_i'(y, v), \qquad i = 1, 2, 3, \tag{2}$$

where the ϕ_i' and ψ_i' are polynomials homogeneous in the y_i of degrees a' and c', and in the v_i of degrees b' and d', respectively, and

$$\rho'' y_i = \phi_i(\phi', \psi'), \qquad \sigma'' v_i = \psi_i(\phi', \psi').$$

This transformation is a line element transformation† if each of the equations

$$\sum u_i x_i = 0, \qquad \sum v_i y_i = 0$$

implies the other. An element transformation is a birational contact transformation if, in addition to the above requirements, each of the two systems of equations

$$\sum x_i du_i = \sum u_i dx_i = 0$$

and

$$\sum y_i dv_i = \sum v_i dy_i = 0$$

implies the other.‡

Simple examples of birational contact transformations in the plane

* Presented to the Society, February 26, 1938.

† A line element transformation is not a contact transformation unless it preserves unions.

‡ The representation of line elements by means of six coordinates x_i, u_i is due to Clebsch who also gave the necessary and sufficient conditions that (1) be a contact transformation. Clebsch-Lindemann-Benoist, *Leçons sur la Géométrie*, Paris, 1883, vol. 3, p. 463.

are furnished by enlarged Cremona point transformations, Cremona line transformations, polar reciprocation, and combinations of these. That birational contact transformations, other than these types, exist was demonstrated by Fano who gave some examples.* Groups of birational contact transformations were first investigated by Autonne, who determined the finite linear and finite quadratic groups of birational transformations in the plane.†

It is the purpose of this paper to construct and study an infinite mixed group of birational contact transformations in S_2. It will then be shown that an analogous group exists in S_n.

2. **The abelian group of birational contact transformations.** We begin with the consideration of a certain directrix equation

$$\sum (y_i/x_i)^{1/k} = 0, \tag{3}$$

k an integer, ($|k| > 1$), which, it will be shown, determines a birational contact transformation belonging to a group of such transformations. In §4 these transformations will be generalized to r-space, and in turn the directrix equations of the generalized transformations will be derived.

By virtue of (3) there corresponds to every point y_i in the yv-plane a triangular-symmetric curve in the xu-plane.‡ Likewise, to every point x_i in the xu-plane corresponds a triangular-symmetric curve in the yv-plane. To a generic curve in either plane corresponds the envelope of a family of triangular-symmetric curves in the other plane.

Differentiating (3), first under the assumption that the y_i are constants and second that the x_i are constants, we obtain

$$\sum_{i=1}^{3} y_i^{1/k} x_i^{-(1+k)/k} dx_i = 0, \tag{4}$$

and

$$\sum x_i^{-1/k} y_i^{(1-k)/k} dy_i = 0. \tag{5}$$

From (3) and (4) it follows that

* G. Fano, *Trasformazione di contatto birazionali del piano*, Rendiconti delle Reale Accademia Nazionale dei Lincei, (6), vol. 8 (1928), p. 445.

† L. Autonne, (1) *Groupes d'ordre fini des substitutions linéaires de contact*, Journal de Mathématique, (4), vol. 3 (1887); (2) *Groupes d'ordre fini contenus dans le groupe quadratique crémonien*, ibid., vol. 4 (1888), pp. 177, 407; (3) Autonne's investigations regarding groups in S_3 and S_n which appear in Annales de l'Université de Lyon, Science et Médecine, vol. 16, 1905.

‡ Triangular-symmetric curves are those of the type $(x_1/a_1)^m+(x_2/a_2)^m+(x_3/a_3)^m=0$. See, for instance, G. Loria, *Spezielle algebraische und transzendente ebene Kurven*, 1910, vol. 1, p. 341.

$$\begin{aligned} \rho y_1 &= x_1^{k+1}(x_2dx_3 - x_3dx_2)^k, \\ \rho y_2 &= x_2^{k+1}(x_3dx_1 - x_1dx_3)^k, \\ \rho y_3 &= x_3^{k+1}(x_1dx_2 - x_2dx_1)^k; \end{aligned} \tag{6}$$

and from (3) and (5) that

$$\begin{aligned} \rho' x_1 &= y_1^{-(k-1)}(y_2dy_3 - y_3dy_2)^{-k}, \\ \rho' x_2 &= y_2^{-(k-1)}(y_3dy_1 - y_1dy_3)^{-k}, \\ \rho' x_3 &= y_3^{-(k-1)}(y_1dy_2 - y_2dy_1)^{-k}. \end{aligned} \tag{7}$$

Since

$$\begin{aligned} u_1 : u_2 : u_3 &= (xdx)_1 : (xdx)_2 : (xdx)_3, \\ v_1 : v_2 : v_3 &= (ydy)_1 : (ydy)_2 : (ydy)_3, \end{aligned}$$

we find, by substituting in (6) and (7) and then solving for v_i and u_i, that the equations of the birational contact transformation T_k and of its inverse T_k^{-1} are

$$T_k: \quad \begin{aligned} &\rho y_i = x_i^{k+1}u_i^k, && i = 1, 2, 3, \\ &\sigma v_1 = (x_2x_3)^k(u_2u_3)^{k-1}, \quad \sigma v_2 = (x_3x_1)^k(u_3u_1)^{k-1}, \\ &\sigma v_3 = (x_1x_2)^k(u_1u_2)^{k-1}, \end{aligned}$$

and

$$T_k^{-1}: \quad \begin{aligned} &\rho' x_1 = (y_2y_3)^{k-1}(v_2v_3)^k, \quad \rho' x_2 = (y_3y_1)^{k-1}(v_3v_1)^k, \\ &\rho' x_3 = (y_1y_2)^{k-1}(v_1v_2)^k, \\ &\sigma' u_i = v_i^{k+1}y_i^k, && i = 1, 2, 3. \end{aligned}$$

From the equations of T_k it can be easily verified that

$$\begin{aligned} \rho\sigma \sum v_iy_i &= (x_1x_2x_3)^k(u_1u_2u_3)^{k-1} \sum u_ix_i, \\ \rho\sigma \sum v_idy_i &= (x_1x_2x_3)^k(u_1u_2u_3)^{k-1}\left[(k+1)\sum u_idx_i + k\sum x_idu_i\right], \\ \rho\sigma \sum y_idv_i &= -(x_1x_2x_3)^k(u_1u_2u_3)^{k-1}\left[k\sum u_idx_i + (k-1)\sum x_idu_i\right]. \end{aligned}$$

Evidently T_k is a birational contact transformation when $k = 0, 1$, so that the restriction previously put on k may be removed.

The triangular-symmetric curve

$$(a_1x_1)^p + (a_2x_2)^p + (a_3x_3)^p = 0 \tag{8}$$

is transformed by T_k into another triangular-symmetric curve, namely,

$$\sum (a_iy_i)^{p/(kp+1)} = 0; \tag{8'}$$

hence T_k leaves a three-parameter family of triangular-symmetric curves invariant. If $p'=p/(kp+1)$, the indices* of (8) and (8′) are connected by the equation

$$\frac{1}{p'}-\frac{1}{p}=k.$$

Since T_0 transforms a generic element into itself, it is an identical transformation and will be represented by the symbol I. We note, furthermore, that

$$T_k^{-1}=T_{-k}, \qquad T_kT_{-k}=I, \qquad T_kT_m=T_mT_k=T_{k+m}.$$

Consequently, the transformations T_k, ($k=\cdots, -2, -1, 0, 1, 2, \cdots$), form a commutative group, G.

THEOREM 1. *The set of ∞^1 transformations T_k (k an integer) forms a discrete abelian group of birational contact transformations that leave a family of ∞^3 triangular-symmetric curves invariant.*

Let the vertices of the fundamental triangle be A_1 (1, 0, 0), A_2 (0, 1, 0), A_3 (0, 0, 1), and let (x, u) be a generic line element. If A, B, C represent lines joining x to A_1, A_2, A_3, respectively, and if D represents u, then the coordinates of these four lines are given by

$$A\colon\ 0, x_3, -x_2, \qquad B\colon\ x_3, 0, -x_1,$$
$$C\colon\ x_2, -x_1, 0, \qquad D\colon\ u_1, u_2, u_3.$$

Evidently

$$x_3C=-x_1A+x_2B, \qquad x_3D=u_2A+u_1B.$$

Therefore

$$R(AB, CD)=-\frac{x_2u_2}{x_1u_1}.$$

If (y, v) is the line element that corresponds to (x, u) by virtue of T_k, and if A', B', C' represent lines yA_1, yA_2, yA_3, respectively, while D' represents v, then the coordinates of A', B', C', D' are as follows:

$$A'\colon\ 0,\ x_3^{k+1}u_3^k,\ -x_2^{k+1}u_2^k,$$
$$B'\colon\ x_3^{k+1}u_3^k,\ 0,\ -x_1^{k+1}u_1^k,$$
$$C'\colon\ x_2^{k+1}u_2^k,\ -x_1^{k+1}u_1^k,\ 0,$$
$$D'\colon\ (x_2x_3)^k(u_2u_3)^{k-1},\ (x_3x_1)^k(u_3u_1)^{k-1},\ (x_1x_2)^k(u_1u_2)^{k-1}.$$

* G. Loria calls p the index of the curve (8); loc. cit., p. 329.

Therefore

$$x_3^{k+1}u_3^k C' = -x_1^{k+1}u_1^k A' + x_2^{k+1}u_2^k B',$$
$$x_3u_3D' = x_1^k u_1^{k-1}A' + x_2^k u_2^{k-1}B'.$$

Consequently,

$$R(A'B', C'D') = -\frac{x_2u_2}{x_1u_1} = R(AB, CD).$$

In the same way it can be shown that the cross ratio that x forms with the points at which u intersects the sides of the fundamental triangle is equal to the cross ratio that the point y of the corresponding line element forms with the points at which v intersects the sides of the fundamental triangle.

THEOREM 2. *Every transformation in the group G sets up a one-to-one correspondence between the line elements of the plane in such manner that, if* (x, u) *and* (y, v) *are corresponding elements,* (1) *the cross ratio that u forms with the lines joining x to the vertices of the fundamental triangle equals the cross ratio that v forms with the lines joining y to the vertices of the fundamental triangle, and dually,* (2) *the cross ratio that x forms with the points at which u intersects the sides of the fundamental triangle equals the cross ratio that y forms with the points at which v intersects the sides of the fundamental triangle.*

The W curves

$$x_1^{a_1}x_2^{a_2}x_3^{a_3} = C, \qquad a_1 + a_2 + a_3 = 0, \tag{9}$$

may be regarded as limits of the triangular-symmetric curves

$$a_1x_1^m + a_2x_2^m + a_3x_3^m = 0, \qquad a_1 + a_2 + a_3 = 0, \qquad a_i \text{ real},$$

when $m \to 0$.* Moreover, each of these W curves has the property that the cross ratio of any one of its points and the three points of intersection of the tangent at the point with the sides of the fundamental triangle is constant for the curve; and dually, the cross ratio of any one of its tangents and the three lines joining the point of tangency to the vertices of the fundamental triangle is constant for the curve.†

* Cesàro-Kowalewski, *Vorlesungen über natürliche Geometrie*, Leipzig, 1901, p. 130.

† The cross ratio associated with any W curve is independent of C, its value being $-a_i/a_j$, $(i \neq j)$, where the subscripts depend on the sequence of the elements forming the cross ratio. To a given cross ratio correspond ∞^1 curves of the family (9), so that (9) constitutes a two-parameter family. It should be noted that only the ratios of the a_i and not the a_i themselves are significant as parameters. See F. Klein, *Vorlesungen über Höhere Geometrie*, Berlin, 1926, p. 170.

Consequently we can add the following corollary to Theorem 2:

COROLLARY. *The birational contact transformations of group G transform each member of a two-parameter family of W curves into another member of the same family possessing the same characteristic cross ratio.*

3. **The ∞^1 involutory birational contact transformations.** The directrix equation

$$\sum_{i=1}^{3} (y_i/u_i)^{1/k} = 0$$

determines a birational contact transformation S_k the equations of which are obtained in the same manner as were the equations of T_k:

$$S_k: \quad \begin{aligned} \rho y_i &= x_i^k u_i^{k+1}, \\ \sigma v_1 &= (x_2x_3)^{k-1}(u_2u_3)^k, \qquad \sigma v_2 = (x_3x_1)^{k-1}(u_3u_1)^k, \\ \sigma v_3 &= (x_1x_2)^{k-1}(u_1u_2)^k, \qquad\qquad i = 1, 2, 3, \end{aligned}$$

and

$$S_k^{-1}: \quad \begin{aligned} \rho' x_i &= y_i^k v_i^{k+1}, \\ \sigma' u_1 &= (y_2y_3)^{k-1}(v_2v_3)^k, \qquad \sigma' u_2 = (y_3y_1)^{k-1}(v_3v_1)^k, \\ \sigma' u_3 &= (y_1y_2)^{k-1}(v_1v_2)^k, \qquad\qquad i = 1, 2, 3. \end{aligned}$$

From the equations of S_k it follows that

$$\rho\sigma \sum y_iv_i = (x_1x_2x_3)^{k-1}(u_1u_2u_3)^k \sum x_iu_i,$$

and

$$\begin{aligned} \rho\sigma \sum y_i dv_i &= -(x_1x_2x_3)^{k-1}(u_1u_2u_3)^k[(k-1)\sum u_i dx_i + k\sum x_i du_i], \\ \rho\sigma \sum v_i dy_i &= (x_1x_2x_3)^{k-1}(u_1u_2u_3)^k[k\sum u_i dx_i + (k+1)\sum x_i du_i]. \end{aligned}$$

It is apparent that S_k is an involutory birational contact transformation. By means of S_k the ∞^3 triangular-symmetric curves (8) are transformed into

$$\sum_{i=1}^{3} (y_i/a_i)^{p/(pk+p-1)} = 0. \tag{8''}$$

If the index of (8'') is represented by p', it is seen that the indices of a triangular-symmetric curve and its transform by S_k are connected by the equation

$$\frac{1}{p} + \frac{1}{p'} = k + 1.$$

Four triangular-symmetric curves are left invariant by S_k, namely,

$$\sum_{i=1}^{3} \pm (x_i)^{2/(k+1)} = 0.$$

Regard S_k as a substitution, and let S_kS_n represent the effect of substitution S_k followed by S_n. Let P represent the enlarged polarity $\rho y_i = u_i$, $\sigma v_i = x_i$. Then it easy to show that

$$\begin{aligned} &S_kS_n = T_{k-n}, && S_k^2 = I, && S_0 = P, \\ &S_nT_k = S_{n-k}, && T_kS_n = T_nS_k = S_{k+n}, \\ &S_kT_k = P, && S_kP = T_k, && PS_k = T_{-k}, \\ &PS_kP = S_{-k}, && S_nT_kS_n = T_{-k}. \end{aligned} \tag{10}$$

The ∞^1 birational contact transformations S_k, (k an integer), do not form a group, but the aggregate set of the S_k and the T_n form a mixed group Γ of which G is evidently an invariant subgroup.

THEOREM 3. *The set of ∞^1 involutory birational contact transformations S_k and the set of ∞^1 birational contact transformations T_n together form a mixed group Γ which contains G as an invariant subgroup.*

For any integer c the set of transformations T_{kc} obviously forms a group G_1, a subgroup of G. Moreover, the set of transformations in G_1 and the set S_{kc} together form a group Γ_1, a subgroup of Γ. If we set up a one-to-one correspondence between the members of Γ_1 and Γ, such that T_{kc} and T_k correspond to one another and S_{kc} and S_k also correspond to one another for all integral values of k, it is seen, in view of the relationships given in (10), that Γ_1 and Γ are isomorphic and so are G_1 and G.

If (x, u) and (y, v) are a pair of corresponding line elements with respect to S_k, it can be shown, as it was for T_k, that the cross ratio formed by u and the lines joining x to the vertices of the fundamental triangle is equal to the cross ratio formed by v and the three lines joining y to the vertices of the fundamental triangle, and a similar result holds in the dual case.

THEOREM 4. *Every transformation in Γ preserves the cross ratio formed by the line of a generic line element and the joins of its point with the vertices of the fundamental triangle; and dually, every transformation in Γ preserves the cross ratio formed by the point of a generic line element and the intersections of its line with the sides of the fundamental triangle.*

It was shown above that the T_k leave the two-parameter family of

W curves (9) invariant. For the same reason the S_k leave the same family invariant.* Consequently, we add the following corollary:

COROLLARY. *The transformations in* Γ *leave a two-parameter family of* W *curves invariant.*

It is interesting to observe that S_n and T_n are determined by other directrix equations in addition to those indicated above.† For instance, the directrix equations

$$\sum (y_i x_i)^{1/(n+1)} = 0, \quad \sum (v_i u_i)^{1/(1-n)} = 0, \quad \sum (v_i/x_i)^{-1/n} = 0$$

determine S_n, whereas

$$\sum (y_i u_i)^{1/(n+1)} = 0, \quad \sum (v_i x_i)^{1/(1-n)} = 0, \quad \sum (v_i/u_i)^{-1/n} = 0$$

determine T_n.

4. **Extension to higher space.** The transformations T_n and S_n have analogs in higher spaces. Thus in r-space, the analog of T_n is given by

$$\bar{T}_n: \quad \rho y_i = x_i^{n+1} u_i^n, \quad \sigma v_i = \frac{1}{x_i^n u_i - 1}, \quad i = 1, 2, \cdots, r+1, \tag{11}$$

and the analog of S_n is given by

$$\bar{S}_n: \quad \rho y_i = x_i^n u_i^{n+1}, \quad \sigma v_i = \frac{1}{x_i^{n-1} u_i^n}, \quad i = 1, 2, \cdots, r+1. \tag{12}$$

From the first $r+1$ equations in (11) it follows that $\rho y_i/x_i = x_i^n u_i^n$. Therefore, using the relationship $\sum u_i x_i = 0$, we obtain a directrix equation of $\bar{T}_n$ in the form

$$\sum_{i=1}^{r+1} (y_i/x_i)^{1/n} = 0.$$

From the second set of $r+1$ equations in (11) the directrix equation of $\bar{T}_n$ in line coordinates is found to be

$$\sum_{i=1}^{r+1} (u_i/v_i)^{1/n} = 0.$$

Directrix equations of $\bar{S}_n$ in mixed point and line coordinates, found

* A detailed proof appears below in connection with a discussion of these transformations in r-space.

† Autonne, Journal de Mathématique, vol. 4 (1888), shows that every contact transformation in mixed coordinates is determined by four directrix equations (primordiales), namely, point-punctual, point-linear, line-punctual, line-linear.

in the same manner, are

$$\sum_{i=1}^{r+1} (y_i/u_i)^{1/n} = 0$$

and

$$\sum_{i=1}^{r+1} (x_i/v_i)^{1/n} = 0.$$

The directrix equations of each of these transformations can be expressed in two other forms, as was indicated above in the planar case.

The aggregate of birational contact transformations, $\overline{T}_n$ and $\overline{S}_n$ (n an integer), can be shown to constitute a group $\overline{\Gamma}$ that has an invariant abelian subgroup $\overline{G}$ composed of the transformations $\overline{T}_n$. The transformation $\overline{S}_n$ is involutory, interchanges the ∞^{r+1} varieties in (13), replacing

$$\sum_{i=1}^{r+1} (a_i x_i)^{1/p} = 0 \tag{13}$$

by

$$\sum (y_i/a_i)^{p/(pn+p-1)} = 0,$$

and leaves each variety in the set

$$\sum \pm (x_i)^{2/(n+1)} = 0$$

invariant. On the other hand, $\overline{T}_n$ transforms the family (13) into itself, replacing (13) by

$$\sum (a_i y_i)^{p/(np+1)} = 0.$$

Because these statements can be proved in the same way that their analogues regarding S_n and T_n in the plane were proved, no verification is offered.

Let us consider the ∞^r varieties

$$F \equiv \prod_{i=1}^{r+1} y_i^{a_i} = K, \qquad \sum_{1}^{r+1} a_i = 0, \tag{14}$$

K constant. This family is evidently a generalization in r-space of the W curves discussed above.* It will be proved that this family is left invariant by the transformations in $\overline{\Gamma}$. First we show that $\overline{T}$ leaves (14) invariant.

* Surfaces of this type in three-space were investigated by Lie and Klein, Comptes Rendus de l'Académie des Sciences, 1870.

Let v_i, $(i=1, 2, \cdots, r+1)$, be the coordinates of the hyperplane tangent to F at y_i; then

$$\alpha v_i = \frac{\partial F}{\partial y_i} = \frac{a_i K}{y_i}, \tag{15}$$

where α is a constant of proportionality. Eliminating the v_i from (15) and from

$$\sigma v_i = \frac{1}{x_i^n u_i^{n-1}},$$

we obtain

$$\alpha' y_i = a_i x_i^n u_i^{n-1} K.$$

If we let the new constant of proportionality absorb the constant K, and if we drop the prime, this equation becomes

$$\alpha y_i = a_i x_i^n u_i^{n-1}. \tag{16}$$

The elimination of the u_i from (16) and $\rho y_i = x_i^{n+1} u_i^n$ results in

$$\rho' y_i = a_i^n x_i. \tag{17}$$

Substituting from (17) in (14) we obtain the transform of (14), namely

$$\prod_1^{r+1} a_i^{na_i} \prod_1^{r+1} x_i^{a_i} = K, \qquad \sum_1^{r+1} a_i = 0.$$

In the same way it can be proved that $\bar{S}_n$ transforms (14) into

$$K \prod_1^{r+1} x_i^{a_i} = \prod_1^{r+1} a_i^{a_i(n+1)}, \qquad \sum_1^{r+1} a_i = 0,$$

so that all the members of $\bar{\Gamma}$ transform this family into itself.*

NEW YORK CITY

* Continuous (Lie) groups in which the discrete groups Γ and $\bar{\Gamma}$ are immersed will be considered in a subsequent paper.

POLYGONAL VARIATIONS*

ALINE H. FRINK AND ORRIN FRINK

So-called direct methods in the calculus of variations, such as those of Tonelli involving lower semi-continuity, sometimes insure the existence of an arc minimizing an integral $J=\int f dx$ in cases where many of the partial derivatives of the integrand function f which occur in the usual theory do not exist. Examples are integrals of the form $\int g(x, y)(1+y'^2)^{1/2}dx$, or in the parametric case $\int g(x, y)(x'^2+y'^2)^{1/2}dt$, where $g(x, y)$ is merely continuous and positive. When a minimizing arc is known to exist, necessary conditions assume greater importance. It seems desirable, therefore, to have methods of deriving the familiar necessary conditions of Weierstrass, Euler, and Legendre, while making as few assumptions as possible concerning the existence of partial derivatives of the integrand f. In this paper it is shown that by using the method of polygonal variations the Weierstrass necessary condition and a generalization of the Euler equation can be derived under the assumption of the existence and continuity of the partial derivative $f_{y'}$ only. A slightly generalized form of the Legendre condition can be proved, with the assumption only of the existence of the generalized second partial derivative $f_{y'y'}$. The method involves giving the dependent variable or variables variations whose graphs are polygonal lines of proper shape, depending on a parameter ϵ, and then evaluating the derivative $J'(\epsilon)$ when $\epsilon=0$. Since the method generalizes easily in the usual way to the case of more than one dependent variable and also to the parametric problem, the discussion will be given here only for the simplest case of a non-parametric problem with one dependent variable.

1. **The Weierstrass necessary condition.** Some of the methods commonly used to derive the Weierstrass necessary condition make use of such complicated notions as fields of extremals and Hilbert's invariant integral, and most of them use the Euler equation, which involves the partial derivative f_y. L. M. Graves however, has given a proof (American Mathematical Monthly, vol. 41 (1934), pp. 502–504) which is independent of the Euler equation and requires only the existence and continuity of $f_{y'}$. The following proof by means of polygonal variations is made under the same weak assumptions.

In all that follows the integrand $f(x, y, y')$ is assumed to be con-

* Presented to the Society, April 15, 1938.

tinuous for x, y in a region R and for all y', and admissible arcs are of class D', lie in R, and pass through the fixed end points (a, c) and (b, d). Since polygonal variations have corners, their use requires the assumption that the integral is minimized with respect to a set of admissible arcs of class D'. In the classical proofs of the necessary conditions, the weaker assumption that comparison arcs are of class C' is all that is needed. This loss of generality could have been avoided by "rounding off the corners" of the polygonal variations. Since this procedure would make the proofs more complicated, and the variations would no longer be polygonal, it is not used here.

THEOREM 1. *If $f_{y'}(x, y, y')$ exists and is continuous for x, y in R and all y', and if $y=y(x)$ makes the integral $J=\int_a^b f(x, y, y')dx$ a strong relative minimum in the class of admissible arcs, then at each element x, y, y' of the minimizing arc and for all k,*

$$E(x, y, y', k) = f(x, y, k) - f(x, y, y') - (k - y')f_{y'}(x, y, y') \geqq 0.$$

PROOF. The proof will be made first for the left-hand end point (a, c). For a given value of k, let $m=k-y'(a)$. The polygonal variation $\eta(x)$ to be used is linear from a to $a+\epsilon$ with slope m, then linear from $a+\epsilon$ to $a+\epsilon+\delta$ with slope $-m\epsilon/\delta$, and identically zero from $a+\epsilon+\delta$ to b, (where δ is to approach zero with ϵ in the manner described below) that is,

$$\begin{aligned} &\eta(x) = m(x - a), \qquad a \leqq x \leqq a + \epsilon, \\ (1)\quad &\eta(x) = m\epsilon(a + \epsilon + \delta - x)/\delta, \qquad a + \epsilon \leqq x \leqq a + \epsilon + \delta, \\ &\eta(x) = 0, \qquad a + \epsilon + \delta \leqq x \leqq b. \end{aligned}$$

Since f is continuous, there exists an $\omega(\epsilon)>\epsilon>0$ such that if $|y_1-y_2| < |m|\epsilon$, then

$$(2) \qquad |f(x, y_1, y') - f(x, y_2, y')| < \omega(\epsilon)$$

for all x, y' on $y=y(x)$, where $\omega(\epsilon)$ approaches zero with ϵ. Now let $\delta=\epsilon/(\omega(\epsilon))^{1/2}$. Since $\omega(\epsilon)>\epsilon$ and $\omega(\epsilon)\to 0$, it is seen that $\delta\to 0$ and $\epsilon/\delta\to 0$. This completes the definition of $\eta(x)$. Since its slope does not approach zero with ϵ, $\eta(x)$ is a strong variation.

Now replace y and y' in J by $y(x)+\eta(x)$ and $y'(x)+\eta'(x)$, and call the result $J(\epsilon)$. Since $y=y(x)$ makes J a strong relative minimum, and ϵ is positive,

$$(3) \qquad J'(0) = \lim_{\epsilon \to 0} [J(\epsilon) - J(0)]/\epsilon \geqq 0$$

provided the limit exists. Adding and subtracting an integral and integrating over subintervals, we have

$$
\begin{aligned}
J'(0) = \lim_{\epsilon \to 0} \Big\{ & \frac{1}{\epsilon} \int_{a}^{a+\epsilon} [f(x, y+\eta, y'+\eta') - f(x, y, y')]dx \\
& + \frac{1}{\epsilon} \int_{a+\epsilon}^{a+\epsilon+\delta} [f(x, y+\eta, y'+\eta') - f(x, y+\eta, y')]dx \\
& + \frac{1}{\epsilon} \int_{a+\epsilon}^{a+\epsilon+\delta} [f(x, y+\eta, y') - f(x, y, y')]dx \Big\} \\
= \lim_{\epsilon \to 0} & (I_1 + I_2 + I_3) \geqq 0.
\end{aligned} \tag{4}
$$

The first law of the mean for integrals and the continuity of f give

$$
\lim_{\epsilon \to 0} I_1 = f[a, y(a), k] - f[a, y(a), y'(a)]. \tag{5}
$$

Likewise the law of the mean for integrals and the mean value theorem for derivatives give

$$
\begin{aligned}
I_2 &= \frac{\delta}{\epsilon} [f(\alpha, y+\eta, y'+\eta') - f(\alpha, y+\eta, y')] \\
&= -mf_{y'}(\alpha, y+\eta, Y'),
\end{aligned} \tag{6}
$$

where y, y', η, η' are evaluated at $x=\alpha$, which is some point of the interval from $a+\epsilon$ to $a+\epsilon+\delta$, and where $y'(\alpha) \leqq Y' \leqq y'(\alpha)+\eta'(\alpha)$. Since $Y' \to y'(a)$,

$$
\lim_{\epsilon \to 0} I_2 = -mf_{y'}[a, y(a), y'(a)]. \tag{7}
$$

The law of the mean for integrals gives

$$
I_3 = \frac{\delta}{\epsilon} \{f[\beta, y(\beta)+\eta(\beta), y'(\beta)] - f[\beta, y(\beta), y'(\beta)]\}, \tag{8}
$$

where $a+\epsilon \leqq \beta \leqq a+\epsilon+\delta$. Then from (2) and the definition of δ,

$$
|I_3| \leqq \frac{\delta}{\epsilon} \omega(\epsilon) = (\omega(\epsilon))^{1/2}. \tag{9}
$$

Hence $I_3 \to 0$; and combining (4), (5), and (7) we obtain

$$
\begin{aligned}
\lim_{\epsilon \to 0} J'(0) = & f[a, y(a), k] - f[a, y(a), y'(a)] \\
& - mf_{y'}[a, y(a), y'(a)] \geqq 0.
\end{aligned} \tag{10}
$$

But since $m = k - y'(a)$, this is the Weierstrass E function, and the Weierstrass necessary condition has been shown to hold at the left-

hand end point. Since $y=y(x)$ also minimizes J on any subinterval, (10) also holds if a is replaced by any value of x between a and b, which proves Theorem 1.

2. **The Euler equation.** The method of polygonal variations can be used to derive a generalization of the Euler equation which does not involve the partial derivative f_y explicitly. However, the method will first be used to derive the ordinary form of the Euler equation, since the proof is very simple and avoids any use of either integration by parts, the fundamental lemma, or the du Bois-Reymond lemma.

It is first assumed that f_y and $f_{y'}$ exist and are continuous, and that $y=y(x)$ makes $J=\int_a^b f(x, y, y')dx$ a weak relative minimum in the class of admissible arcs joining (a, c) and (b, d). For a fixed $\delta>0$, define $\eta(x, \delta)$ as follows:

$$
\begin{aligned}
&\eta(x, \delta) = (x-a)/\delta, && a \leqq x \leqq a+\delta,\\
&\eta(x, \delta) = 1, && a+\delta \leqq x \leqq b-\delta,\\
&\eta(x, \delta) = (b-x)/\delta, && b-\delta \leqq x \leqq b.
\end{aligned} \tag{11}
$$

Replace y and y' in J by $y(x)+\epsilon\eta(x)$ and $y'(x)+\epsilon\eta'(x)$, and call the result $I(\epsilon)$. The polygonal variation $\eta(x)$ is linear on the interval $[a, a+\delta]$ with slope $1/\delta$, constant on the interval $[a+\delta, b-\delta]$, and linear with slope $-1/\delta$ on the interval $[b-\delta, b]$. Since $y=y(x)$ makes J a weak relative minimum, it follows in the usual way that

$$
I'(0) = \int_a^b (\eta f_y + \eta' f_{y'})dx = 0 \tag{12}
$$

for all $\delta>0$. Hence, by integrating over subintervals and taking the limit as $\delta\to 0$,

$$
\begin{aligned}
\lim_{\delta\to 0} \int_a^b (\eta f_y + \eta' f_{y'})dx = \lim_{\delta\to 0} \Big\{ &\int_a^{a+\delta} \frac{x-a}{\delta} f_y dx + \int_{a+\delta}^{b-\delta} 1\cdot f_y dx\\
&+ \int_{b-\delta}^{b} \frac{b-x}{\delta} f_y dx + \int_a^{a+\delta} \frac{1}{\delta} f_{y'} dx + 0 + \int_{b-\delta}^{b} -\frac{1}{\delta} f_{y'} dx \Big\} = 0.
\end{aligned} \tag{13}
$$

The first and third of these integrals approach the limit zero, as may be seen by using the law of the mean for integrals. Likewise the last two integrals approach the limits $f_{y'}(a)$ and $-f_{y'}(b)$, respectively. The limit of the second integral is seen to be $\int_a^b f_y dx$. Hence

$$
\int_a^b f_y dx = f_{y'}(b) - f_{y'}(a). \tag{14}
$$

Since $y=y(x)$ also minimizes J on any subinterval, (14) holds with b

replaced by any value of x from a to b, so that

$$\int_a^x f_y dx = f_{y'}(x) - f_{y'}(a). \tag{15}$$

But this is the Euler equation in integrated form.

3. **Partial variations and the generalized Euler equation.** The partial derivative f_y occurs in the Euler equation (14) in the integral $\int_a^b f_y dx$. This integral can also be obtained in the following way. Give the dependent variable y the (non-admissible) polygonal variation $\eta(x) = \epsilon$, a constant. Then $\eta'(x) = 0$, and the integral J becomes

$$J(\epsilon) = \int_a^b f[x, y(x) + \epsilon, y'(x)]dx.$$

Then, if f_y exists and is continuous, $J'(0) = \int_a^b f_y dx$. Examples show that $J'(0)$ may exist even if f_y does not. This suggests the definition

$$\partial J = J'(0) = \lim_{\epsilon \to 0} \frac{1}{\epsilon} \int_a^b \{f[x, y(x) + \epsilon, y'(x)] - f[x, y(x), y'(x)]\}dx.$$

We shall call ∂J the *partial variation* of J since only y and not y' is varied. The partial variation ∂J is seen to be a generalization of the integral $\int_a^b f_y dx$ which occurs in the Euler equation. The variation $\eta(x) = \epsilon$ is not admissible since it does not vanish at the end points; hence ∂J cannot be expected to vanish. We have rather the following theorem:

THEOREM 2. *If f and $f_{y'}$ are continuous for x, y in R and all y', and if $y = y(x)$ makes $J = \int_a^b f(x, y, y')dx$ a weak relative minimum in the class of admissible functions, and ∂J exists, then $\partial J = f_{y'}(b) - f_{y'}(a)$.*

PROOF. Given ϵ, define the weak polygonal variation $\eta(x)$ as follows:

$$\begin{aligned} \eta(x) &= \epsilon(x - a)/\delta, \qquad a \leqq x \leqq a + \delta, \\ \eta(x) &= \epsilon, \qquad a + \delta \leqq x \leqq b - \delta, \\ \eta(x) &= \epsilon(b - x)/\delta, \qquad b - \delta \leqq x \leqq b. \end{aligned} \tag{16}$$

This variation $\eta(x)$ is linear from a to $a+\delta$ with slope ϵ/δ which is to approach zero with ϵ. From $a+\delta$ to $b-\delta$, $\eta(x)$ is constant, and from $b-\delta$ to b, $\eta(x)$ is linear with a slope which also approaches zero with ϵ. As in the proof of Theorem 1, δ is defined in terms of ϵ as follows. From the continuity of f, there exists $\omega(\epsilon) > |\epsilon|$ such that, if

$|y_1-y_2| \leqq |\epsilon|$,

$$|f(x, y_1, y') - f(x, y_2, y')| < \omega(\epsilon) \tag{17}$$

for x, y' on $y=y(x)$, where $\omega(\epsilon)\to 0$ as $\epsilon\to 0$. Now let $\delta=|\epsilon|/(\omega(\epsilon))^{1/2}$. Note that as $\epsilon\to 0$, $\delta\to 0$, and $\epsilon/\delta\to 0$, also that δ is positive, while ϵ may be positive or negative.

Replace y and y' in J by $y(x)+\eta(x)$ and $y'(x)+\eta'(x)$, and call the result $J(\epsilon)$. Since $y=y(x)$ makes J a weak relative minimum, it follows that

$$V = \lim_{\epsilon\to 0} \frac{1}{\epsilon} \int_a^b [f(x, y+\eta, y'+\eta') - f(x, y, y')]dx = 0$$

provided the limit V exists, since the integral is non-negative for ϵ sufficiently small, while ϵ may be positive or negative. After adding and subtracting two quantities this becomes

$$\begin{aligned} V = \lim_{\epsilon\to 0} \frac{1}{\epsilon} \Big\{ & \int_a^b [f(x, y+\eta, y'+\eta') - f(x, y+\eta, y')]dx \\ & + \int_a^b [f(x, y+\eta, y') - f(x, y+\epsilon, y')]dx \\ & + \int_a^b [f(x, y+\epsilon, y') - f(x, y, y')]dx \Big\} \\ = \lim_{\epsilon\to 0} \frac{1}{\epsilon} \Big\{ & \int_a^b F_1 dx + \int_a^b F_2 dx + \int_a^b F_3 dx \Big\} = 0. \end{aligned} \tag{18}$$

Here F_1, F_2, F_3 have been introduced for convenience. By the law of the mean for integrals and the mean value theorem for derivatives,

$$\int_a^{a+\delta} F_1 dx = \epsilon f_{y'}[\alpha, y(\alpha)+\eta(\alpha), Y'], \tag{19}$$

where $a\leqq\alpha\leqq a+\delta$ and $y'(\alpha)\leqq Y'\leqq y'(\alpha)+\eta'(\alpha)$. Since $Y'\to y'(a)$, and $f_{y'}$ is continuous,

$$\lim_{\epsilon\to 0} \frac{1}{\epsilon} \int_a^{a+\delta} F_1 dx = f_{y'}[a, y(a), y'(a)]. \tag{20}$$

Similarly,

$$\lim_{\epsilon\to 0} \frac{1}{\epsilon} \int_{b-\delta}^{b} F_1 dx = -f_{y'}[b, y(b), y'(b)]. \tag{21}$$

Since F_1 vanishes from $a+\delta$ to $b-\delta$,

$$\lim_{\epsilon\to 0}\frac{1}{\epsilon}\int_a^b F_1 dx = f_{y'}[a, y(a), y'(a)] - f_{y'}[b, y(b), y'(b)]. \tag{22}$$

Consider now $\int_a^b F_2 dx$. Since $|\eta(x)-\epsilon| \leqq |\epsilon|$, by the definition of $\omega(\epsilon)$,

$$|F_2| = |f(x, y+\eta, y') - f(x, y+\epsilon, y')| < \omega(\epsilon). \tag{23}$$

Hence, by the law of the mean,

$$\left|\frac{1}{\epsilon}\int_a^{a+\delta} F_2 dx\right| \leqq \frac{\delta}{|\epsilon|}\omega(\epsilon) = (\omega(\epsilon))^{1/2}. \tag{24}$$

Therefore, since $\omega(\epsilon)\to 0$,

$$\lim_{\epsilon\to 0}\frac{1}{\epsilon}\int_a^{a+\delta} F_2 dx = 0. \tag{25}$$

Similarly,

$$\lim_{\epsilon\to 0}\frac{1}{\epsilon}\int_{b-\delta}^{b} F_2 dx = 0. \tag{26}$$

Since $F_2 = 0$ for $a+\delta \leqq x \leqq b-\delta$, (25) and (26) give

$$\lim_{\epsilon\to 0}\frac{1}{\epsilon}\int_a^b F_2 dx = 0. \tag{27}$$

From the definition of ∂J,

$$\lim_{\epsilon\to 0}\frac{1}{\epsilon}\int_a^b F_3 dx = \partial J. \tag{28}$$

Combining (22), (27), and (28) we have

$$V = f_{y'}[a, y(a), y'(a)] - f_{y'}[b, y(b), y'(b)] + \partial J = 0, \tag{29}$$

which proves Theorem 2. Note that since $y=y(x)$ also minimizes J over any subinterval, b in (29) can be replaced by any value of x between a and b, provided that ∂J is interpreted to mean the partial variation of J over the interval from a to x.

4. **The Legendre condition.** The Legendre condition is usually derived either as a consequence of the Weierstrass condition or by use of the second variation. The objection to deriving it as a consequence of the Weierstrass condition is that the Legendre condition is necessary for a weak relative minimum while the Weierstrass condition is not. To derive the Legendre condition from the second variation re-

quires the existence and continuity of numerous partial derivatives of f. Here a generalization of the Legendre condition is derived by the use of a weak polygonal variation, the proof requiring the existence only of the single generalized partial derivative $f_{y'y'}$.

THEOREM 3. *If $f(x, y, y')$ is continuous, and if the arc $y=y(x)$ makes $J=\int_a^b f dx$ a weak relative minimum in the set of arcs of class D' joining (a, c) and (b, d), and if*

$$L = \lim_{m\to 0} \frac{1}{m^2}\{f[x, y(x), y'(x)+m] - 2f[x, y(x), y'(x)] + f[x, y(x), y'(x)-m]\}$$

exists, where x is any point of the interval $[a, b]$, then $L \geqq 0$.

The proof will be given first for the case where x is the left-hand end point a. The weak polygonal variation $\eta(x)$ to be used is linear from a to $a+\epsilon$ with slope m which is to approach zero. It is linear from $a+\epsilon$ to $a+2\epsilon$ with slope $-m$, and identically zero from $a+2\epsilon$ to b; that is,

$$\begin{aligned} &\eta(x) = m(x-a), && a \leqq x \leqq a+\epsilon, \\ &\eta(x) = m(a+2\epsilon-x), && a+\epsilon \leqq x \leqq a+2\epsilon, \\ &\eta(x) = 0, && a+2\epsilon \leqq x \leqq b, \end{aligned} \tag{30}$$

where ϵ is to be defined in terms of m as follows: From the continuity of f, $y(x)$, and $y'(x)$, for every value of m there exists an $\epsilon(m)>0$ such that $\epsilon(m)<|m|$ and such that if $|x-a|<2\epsilon$ and $|\delta|<\epsilon$, the following three inequalities hold:

$$\begin{aligned} &|f[x, y(x)+\delta, y'(x)+m] - f[a, y(a), y'(a)+m]| < m^4, \\ &|f[x, y(x)+\delta, y'(x)-m] - f[a, y(a), y'(a)-m]| < m^4, \\ &|f[a, y(a), y'(a)] - f[x, y(x), y'(x)]| < m^4. \end{aligned} \tag{31}$$

This completes the definition of $\eta(x)$. Since $y=y(x)$ makes J a weak relative minimum, it follows that

$$I = \lim_{m\to 0} \frac{1}{\epsilon m^2}\int_a^b \{f[x, y(x)+\eta(x), y'(x)+\eta'(x)] - f[x, y(x), y'(x)]\}dx \geqq 0, \tag{32}$$

provided the limit exists, since the integral is non-negative for m sufficiently small, and $\epsilon>0$. On the interval from $a+2\epsilon$ to b the integrand in (32) is zero; hence by integrating over subintervals and applying the law of the mean, we can write (32) as

$$(33)\quad I = \lim_{m\to 0} \frac{1}{m^2} \{f[\alpha, y(\alpha) + \eta(\alpha), y'(\alpha) + m] - f[\alpha, y(\alpha), y'(\alpha)] + f[\beta, y(\beta) + \eta(\beta), y'(\beta) - m] - f[\beta, y(\beta), y'(\beta)]\},$$

where $a \leqq \alpha \leqq a+\epsilon$, and $a+\epsilon \leqq \beta \leqq a+2\epsilon$, and use has been made of the fact that $\eta'(\alpha) = m$ and $\eta'(\beta) = -m$ from (30).

After adding and subtracting a quantity, (33) becomes

$$(34)\quad \begin{aligned} I = \lim_{m\to 0} \frac{1}{m^2} \{&f[a, y(a), y'(a) + m] - 2f[a, y(a), y'(a)] \\ &+ f[a, y(a), y'(a) - m]\} + \lim_{m\to 0} \frac{1}{m^2} \{f[\alpha, y(\alpha) + \eta(\alpha), y'(\alpha) + m] \\ &- f[a, y(a), y'(a) + m] + f[a, y(a), y'(a)] - f[\alpha, y(\alpha), y'(\alpha)] \\ &+ f[\beta, y(\beta) + \eta(\beta), y'(\beta) - m] - f[a, y(a), y'(a) - m] \\ &+ f[a, y(a), y'(a)] - f[\beta, y(\beta), y'(\beta)]\} \\ = \lim_{m\to 0} \frac{1}{m^2} B_1 &+ \lim_{m\to 0} \frac{1}{m^2} B_2 \geqq 0. \end{aligned}$$

The two expressions in braces in (34) are here designated B_1, B_2. Now apply (31) to B_2, replacing δ by $\eta(\alpha)$ and x by α in the first line of (31), replacing δ by $\eta(\beta)$ and x by β in the second line of (31), and then replacing x first by α and then by β in the third line of (31). Then

$$(35)\quad |B_2| < 4m^4,$$

since $|\eta(x)| < \epsilon$ for $|m| < 1$ from (30). Hence

$$(36)\quad \lim_{m\to 0} \frac{1}{m_2} B_2 = 0.$$

Therefore, from (34) and the definition of L,

$$(37)\quad L = \lim_{m\to 0} \frac{1}{m_2} B_1 \geqq 0,$$

which proves Theorem 3 when x is the left-hand end point a. Since $y = y(x)$ minimizes J on any subinterval, the same conclusion holds when a is replaced by any value of x between a and b. It is well known that if the ordinary second partial derivative $f_{y'y'}$ exists, it is equal to L; but L may exist even when the first partial derivative $f_{y'}$ does not. Hence Theorem 3 is a generalization of Legendre's condition.

THE PENNSYLVANIA STATE COLLEGE

A PROPERTY OF HARMONIC FUNCTIONS IN THREE VARIABLES*

J. W. GREEN

A harmonic function $u(x, y)$ of the two variables x, y, which is defined in a circle, and which has a normal derivative vanishing on an open set of points on the circle, may be uniquely extended to all points exterior to the circle by assigning the same value to u at any point P' outside the circle as is assumed by u at the point P, image of P' with respect to the circle. The resulting function is harmonic in the entire plane except on the complement of the given open set with respect to the circumference of the circle. In this paper it will be investigated whether or not an analogous result holds for harmonic functions of three variables.

Let r, ϕ, θ be spherical coordinates, and let S be the sphere $r=1$. We consider a function $u(r, \phi, \theta)$ which is harmonic in the interior of the sphere and which together with its normal derivative is continuous on S. We suppose that Ω is a set of points of S, open with respect to S, on which $(\partial u/\partial r)=0$. If M is a point interior to the sphere, then M' is the image of M with respect to S; that is, M' is on the ray OM, where O represents the origin, and $OM \cdot OM' = 1$. On occasion the notations $u(r)$ and $u(r')$ will be used; they refer actually to $u(r, \phi, \theta)$ and $u(r', \phi, \theta)$, where (r, ϕ, θ) and (r', ϕ, θ) are images with respect to S, so that $rr' = 1$. The variables ϕ and θ are omitted when there is no cause for confusion. The symbol Ω is used to denote either the set of points on the sphere or the set of values of ϕ and θ corresponding to these points; in each case the meaning will be clear. If a point $M(r, \phi, \theta)$ is under discussion, Q will denote the point on S on the ray OM.

Let Σ be an open domain containing S and its interior. We have the following theorem:

THEOREM 1. *A necessary and sufficient condition that there exist a unique analytic extension of u across Ω into the portion of Σ exterior to S is that on Ω*

$$\int_0^1 u(r)dr = \text{constant}.$$

Suppose that u is extensible across Ω as described. Let $v(r')$ be the

* Presented to the Society, November 27, 1937.

function defined in the part of Σ outside S which is the harmonic extension of $u(r)$.

LEMMA 1. *On S, the function $w(Q)=v(Q)-u(Q)$ satisfies the equation*

$$\Delta_2 w(Q) = \left(\frac{\partial u}{\partial r}\right)_Q, \tag{1}$$

where Δ_2 is the second differential parameter of Beltrami,

$$\Delta_2 = \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2} + \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right).$$

Consider the function $V(r, \phi, \theta)=r(\partial u/\partial r)$. It is defined and harmonic inside S; in fact,

$$\Delta^2\left(r\frac{\partial u}{\partial r}\right) = \frac{1}{r}\frac{\partial}{\partial r}(r^2\Delta^2 u) = 0.$$

Furthermore V vanishes on Ω. Its harmonic extension across Ω is given by the formula

$$V(M') = -V(M)\cdot OM \qquad \text{or} \qquad V(r') = -rV(r).$$

Inside S, $V(r)=r(\partial u/\partial r)$, and by analytic extension $V(r)$ will be representable in this form everywhere. Outside S then, $V(r')=r'[\partial v(r')/\partial r']$. Equating the two expressions obtained for $V(r')$, we have

$$r'\frac{\partial v(r')}{\partial r'} = V(r') = -rV(r) = -r^2\frac{\partial u(r)}{\partial r},$$

or

$$\frac{\partial v(r')}{\partial r'} = -r^3\frac{\partial u(r)}{\partial r}.$$

Now

$$v(r') = v(Q) + \int_1^{r'}\frac{\partial v}{\partial r'}dr' = v(Q) + \int_1^{r'} -r^3\frac{\partial u(r)}{\partial r}dr'$$

$$= v(Q) + \int_1^r r\frac{\partial u}{\partial r}dr.$$

On integrating this by parts, we obtain

$$v(r') = v(Q) - u(Q) + ru(r) - \int_1^r u dr. \tag{2}$$

Since v is harmonic, it satisfies Laplace's equation,

$$\Delta^2 v(r') = \frac{1}{r'^2}\left\{\frac{\partial}{\partial r'}\left(r'^2\frac{\partial v}{\partial r'}\right) + \Delta_2 v\right\} = 0. \tag{3}$$

If we put $\partial v(r')/\partial r' = -r^2(\partial u(r)/\partial r)$, since $\partial/\partial r' = (\partial/\partial r)(dr/dr') = -r^2(\partial/\partial r)$, we see that (3) reduces to

$$r'^2\Delta^2 v(r') = r^2\frac{\partial}{\partial r}\left(r\frac{\partial u}{\partial r}\right) + \Delta_2 v = 0. \tag{3'}$$

From (2), we obtain $\Delta_2 v(r') = \Delta_2\{v(Q) - u(Q)\} + r\Delta_2 u(r) - \int_1^r \Delta_2 u dr$. Since u is harmonic,

$$\Delta_2 u = -\frac{\partial}{\partial r}\left(r^2\frac{\partial u}{\partial r}\right).$$

Integrating $\Delta_2 u$ from 1 to r, we have

$$\int_1^r \Delta_2 u dr = -\int_1^r \frac{\partial}{\partial r}\left(r^2\frac{\partial u}{\partial r}\right)dr = -r^2\frac{\partial u}{\partial r} + \left(\frac{\partial u}{\partial r}\right)_Q.$$

By employing these values of $\Delta_2 u$ and $\int_1^r \Delta_2 u dr$ we reduce $\Delta_2 v(r')$ to the form

$$\begin{aligned}\Delta_2 v(r') &= \Delta_2\{v(Q) - u(Q)\} - r\frac{\partial}{\partial r}\left(r^2\frac{\partial u}{\partial r}\right) + r^2\frac{\partial u}{\partial r} - \left(\frac{\partial u}{\partial r}\right)_Q \\ &= \Delta_2\{v(Q) - u(Q)\} - r^2\frac{\partial}{\partial r}\left(r\frac{\partial u}{\partial r}\right) - \left(\frac{\partial u}{\partial r}\right)_Q.\end{aligned}$$

If we insert this last value for $\Delta_2 v(r')$ into (3'), it results immediately that

$$r'^2\Delta^2 v(r') = \Delta_2\{v(Q) - u(Q)\} - \left(\frac{\partial u}{\partial r}\right)_Q = 0. \tag{4}$$

If we put $w(Q) = v(Q) - u(Q)$, equation (4) implies that $\Delta_2 w = (\partial u/\partial r)_Q$, and the proof of the lemma is completed.

Let $G(Q)$ and $H(Q)$ be any functions on S with continuous second derivatives, and let D be a domain on S bounded by a curve C with continuous normal. We have the relation*

$$\int_C\left(H\frac{\partial G}{\partial n} - G\frac{\partial H}{\partial n}\right)ds - \iint_D (H\Delta_2 G - G\Delta_2 H)dS = 0. \tag{5}$$

* See for example Hadamard, *Leçons sur la Propagation des Ondes*, Paris, 1903, p. 49.

The differentiation $\partial/\partial n$ is with respect to the outer normal to C. Let P be a point of S, and let C be a small circle about P. Let D be the larger of the two domains determined. For H take the function $w(Q)$ as defined in Lemma 1. Then $\Delta_2 w(Q)$ is $(\partial u/\partial r)_Q$, and we denote this latter function by $F(Q)$. For $G(Q)$ we take the function $\log \sin \psi/2$ where $\psi = \sphericalangle(OP, OQ)$. By differentiation, $\Delta_2 \log \sin \psi/2$ is found to be $-1/2$, and the second integral of (5) reduces to

$$-\frac{1}{2}\iint_D w(Q)dS_Q - \iint_D F(Q) \log \sin \frac{\psi}{2}\, dS_Q .$$

If the radius of C is allowed to approach zero, it is shown without difficulty that the first integral of (5) approaches $-2\pi w(P)$. We then have

$$(6) \qquad w(P) = \frac{1}{4\pi}\iint_S w(Q)dS_Q + \frac{1}{2\pi}\iint_S F(Q) \log \sin \frac{\psi}{2}\, dS_Q .$$

The first term of the right-hand member of (6) does not depend on P. On Ω, we have $w(P) = 0$, since there $v(P) = u(P)$. Consequently, we obtain

$$(7) \qquad \iint_S F(Q) \log \sin \frac{\psi}{2}\, dS_Q = \text{constant}$$

for P on Ω.

LEMMA 2. *If a continuous function $F(Q)$ represents the values assumed by the normal derivative of a harmonic function u on S, then*

$$(8) \qquad \int_0^1 u(r)dr = -\frac{1}{2\pi}\iint_S F(Q) \log \sin \frac{\psi}{2}\, dS_Q + u(0).$$

Here the integration is performed along the ray OP, and as before, $\psi = \sphericalangle(OP, OQ)$.

Poisson's formula gives the value of a harmonic function $h(P)$ at a point P interior to S in terms of its values $h(Q)$ on S,

$$(9) \qquad h(P) = \frac{1}{4\pi}\iint_S \frac{h(Q)(1 - r^2)}{(1 + r^2 - 2r \cos \psi)^{3/2}}\, dS_Q .$$

In particular, apply the formula (9) to the harmonic function $r(\partial u/\partial r)$, which assumes the values $(\partial u/\partial r)_Q = F(Q)$ on S. We have

$$r\frac{\partial u}{\partial r} = \frac{1}{4\pi}\iint_S \frac{F(Q)(1 - r^2)}{(1 + r^2 - 2r \cos \psi)^{3/2}}\, dS_Q ,$$

or

$$\frac{\partial u}{\partial r} = \frac{1}{4\pi}\iint_S \frac{F(Q)(1-r^2)}{r(1+r^2-2r\cos\psi)^{3/2}}\,dS_Q.$$

If r and $\epsilon>0$ represent two points inside S on the same ray through O, then

$$u(r) - u(\epsilon) = \frac{1}{4\pi}\int_\epsilon^r dt \iint_S \frac{F(Q)(1-t^2)}{t(1+t^2-2t\cos\psi)^{3/2}}\,dS_Q$$

$$= \frac{1}{4\pi}\iint_S F(Q)dS_Q \int_\epsilon^r \frac{1-t^2}{t(1+t^2-2t\cos\psi)^{3/2}}\,dt.$$

Since u is harmonic, $\iint_S F(Q)dS_Q=0$, and the above is the same as

$$u(r) - u(\epsilon)$$

$$= \frac{1}{4\pi}\iint_S F(Q)dS_Q\left\{\int_\epsilon^r \left(\frac{1-t^2}{t(1+t^2-2t\cos\psi)^{3/2}} - \frac{1}{t}\right)dt\right\}.$$

Evaluating the inner integral, we find

$$u(r) - u(\epsilon) = \frac{1}{4\pi}\iint_S F(Q)dS_Q\left[\frac{2}{T} + \log\frac{t+T-1}{t(t+T+1)}\right]_\epsilon^r,$$

where $T=(1+t^2-2t\cos\psi)^{1/2}$. Let ϵ approach zero. The bracketed expression in the integral, evaluated for $t=\epsilon$, approaches $2+\log(1-\cos\psi)-\log 2$, and if this value is inserted, we obtain

$$u(r) - u(0)$$

$$(10)\qquad = \frac{1}{4\pi}\iint_S F(Q)dS_Q\left\{\frac{2}{R} + \log\frac{r+R-1}{r(r+R+1)} - \log(1-\cos\psi)\right\},$$

where $R=(1+r^2-2r\cos\psi)^{1/2}$. By combining the last two terms of the right-hand member of (10) we can reduce it to the simpler form

$$(10')\qquad u(r) - u(0) = \frac{1}{4\pi}\iint_S F(Q)dS_Q\left\{\frac{2}{R} - \log(1+R-r\cos\psi)\right\}.$$

We are now in a position to calculate $\int_0^1 u\,dr$. From (10),

$$\int_0^1 u(r)dr = \frac{1}{4\pi}\iint_S F(Q)dS_Q\int_0^1 \left\{\frac{2}{R} + \log\frac{r+R-1}{r(r+R+1)}\right.$$

$$\left. - \log(1-\cos\psi)\right\}dr + u(0).$$

The three definite integrals involved in the right-hand member have the following values:

$$\int_0^1 \frac{2dr}{(1+r^2-2r\cos\psi)^{1/2}} = 2\log\left(1+\sin\frac{\psi}{2}\right) - 2\log\sin\frac{\psi}{2},$$

$$\int_0^1 \log\frac{r+R-1}{r(r+R+1)}dr = 2\log\sin\frac{\psi}{2} - 2\log\left(1+\sin\frac{\psi}{2}\right) + 1,$$

$$\int_0^1 \log(1-\cos\psi)dr = \log(1-\cos\psi) = 2\log\sin\frac{\psi}{2} + \log 2.$$

Substituting for the three definite integrals their values as just computed, we have

$$\int_0^1 u dr = -\frac{1}{2\pi}\int\int_S F(Q)\log\sin\frac{\psi}{2}\,dS_Q + u(0),$$

and the lemma is proved.

From (7) and (8) it follows that $\int_0^1 u dr$ is constant on Ω, and the necessity of the condition of the theorem is proved.

To prove the sufficiency, suppose that $\int_0^1 u dr = c$ on Ω. Then for Q on S, we take for $w(Q) = v(Q) - u(Q)$ the function $c - \int_0^1 u dr$. The function $v(r')$ as given by (2) will actually represent the harmonic extension of u across Ω into the entire exterior of the sphere. For (2) is then

$$\begin{aligned} v(r') &= w(Q) + ru(r) - \int_1^r u dr = c - \int_0^1 u dr + ru(r) - \int_1^r u dr \\ &= c + ru(r) - \int_0^r u dr. \end{aligned} \tag{11}$$

This function $v(r')$ has the following three properties:

(a) $v(Q) = c + u(Q) - \int_0^1 u dr = u(Q)$ on Ω;

(b) $(\partial v/\partial r')_{r'=1} = [-r^2(r(\partial u/\partial r) + u - u)]_{r=1} = -(\partial u/\partial r) = 0$ on Ω;

(c) from (4),

$$\begin{aligned} \Delta^2 v(r') &= \frac{1}{r'^2}\left\{\Delta_2 w(Q) - \left(\frac{\partial u}{\partial r}\right)_Q\right\} = \frac{1}{r'^2}\left\{-\int_0^1 \Delta_2 u dr - \left(\frac{\partial u}{\partial r}\right)_Q\right\} \\ &= \frac{1}{r'^2}\left\{\int_0^1 \frac{\partial}{\partial r}\left(r^2\frac{\partial u}{\partial r}\right)dr - \left(\frac{\partial u}{\partial r}\right)_Q\right\} = 0 \end{aligned}$$

outside S.

Properties (a), (b), and (c) show clearly that $v(r')$ is the harmonic extension of u across Ω. This completes the proof of the theorem.

It is of interest to note that we can prescribe the values of $\int_0^1 u dr$ on the entire sphere and determine the harmonic function u giving these values. More precisely, we prove the following theorem:

THEOREM 2. *Let $M(P)$ be a function with continuous second derivatives defined on S. There exists a unique function u defined and harmonic inside S and such that $\int_0^1 u dr = M(P)$; and u may be expressed in terms of $M(P)$ by means of an integral formula.*

If a function u does exist as described, then

$$u(0) = \frac{1}{4\pi}\int\int_S u(r)dS_Q,$$

where r is any radius not exceeding one and dS_Q is an element of the area of the unit sphere. Thus

$$u(0) = \int_0^1 u(0)dr = \frac{1}{4\pi}\int\int_S \left\{\int_0^1 u(r)dr\right\} dS_Q = \frac{1}{4\pi}\int\int_S M(Q)dS_Q.$$

From (8) we see that $F(Q)$, the value of the normal derivative of u on S, must satisfy the equation

$$(12) \quad M(P) = -\frac{1}{2\pi}\int\int_S F(Q) \log \sin \frac{\psi}{2}\, dS_Q + \frac{1}{4\pi}\int\int_S M(Q)dS_Q.$$

To determine $F(P)$ perform the operation Δ_2 on both sides of the integral equation (12). The first term of the right-hand member gives a term $-F(P)$ because of the singular integrand $\log \sin \psi/2$,* and in addition a term coming from the differentiation under the integral sign:

$$\Delta_2 M(P) = -F(P) + \frac{1}{4\pi}\int\int_S F(Q)dS_Q.$$

From (12), however, we obtain

$$\int\int_S M(P)dS_P = -\frac{1}{2\pi}\int\int_S F(Q)dS_Q \int\int_S \log \sin \frac{\psi}{2}\, dS_P + \int\int_S M(P)dS_P.$$

* The argument is precisely analogous to that used in proving Poisson's equation for the logarithmic potential; see O. D. Kellogg, *Foundations of Potential Theory*, Berlin, 1929, p. 156. For information about the differential parameter Δ_2, see Darboux, *Leçons sur la Théorie des Surfaces*, Paris, 1914, book 7, chap. 1.

Now $\iint_S \log \sin \psi/2 \, dS_r \neq 0$, for the integrand is of constant sign; therefore $\iint_S F(Q) dS_Q = 0$. Hence

$$F(P) = -\Delta_2 M(P). \tag{13}$$

If (12) has a solution $F(Q)$, that solution must be given by (13). That this $F(Q)$ actually satisfies (12) follows from (6). Using the values of $F(Q)$ obtained from (13) in the formula (10), we determine a harmonic function u whose normal derivative on S is $F(Q)$ and for which

$$\int_0^1 u dr = M(P).$$

In particular we can choose $M(P) = c_1$ on some open set Ω_1, and $M(P) = c_2$ on another open set Ω_2, with $\Omega_1 \cdot \Omega_2 = 0$, and give to M elsewhere any convenient values making M have continuous second derivatives. The harmonic function u determined inside of S by the use of (13) and (10) has the properties:

$$\int_0^1 u dr = c_1 \text{ on } \Omega_1,$$

$$\int_0^1 u dr = c_2 \text{ on } \Omega_2,$$

$$\left(\frac{\partial u}{\partial r}\right) = 0 \text{ on } \Omega_1 \text{ and } \Omega_2.$$

This last statement follows from the fact that $(\partial u/\partial r)_P = -\Delta_2 M(P)$. From (11), u may be extended harmonically across Ω_1 or Ω_2 according to the respective formulas

$$u_1(r') = c_1 + ru(r) - \int_0^r u dr,$$

and

$$u_2(r') = c_2 + ru(r) - \int_0^r u dr.$$

The two extensions u_1 and u_2 differ by $c_1 - c_2$. This function u, then, is an example of a harmonic function whose normal derivative vanishes on the open set $(\Omega_1 + \Omega_2)$, and which may not be extended harmonically across this set into the exterior of S in a unique fashion. If u is continued across Ω_1 into the exterior of S and back across Ω_2 into the interior of S, it suffers a jump of $(c_1 - c_2)$.

The case $\partial u/\partial r = c$, $(c \neq 0)$, on an open set Ω can be treated in the same manner and yields the curious result that in this case u cannot be extended harmonically across Ω into a domain Σ containing the sphere S. For suppose that u could be so extended. Then we would consider

$$V(r) = r\frac{\partial u}{\partial r} - c,$$

a harmonic function vanishing on Ω. Its extension across Ω is

$$V(r') = -rV(r) = -r^2\frac{\partial u}{\partial r} + cr.$$

Inside the sphere, $V = r(\partial u/\partial r) - c$. By its analyticity V will be expressible everywhere in this form; hence

$$V(r') = r'\frac{\partial v}{\partial r'} - c,$$

where v is the extension of u into the part of Σ exterior to S. If we equate the two values obtained for $V(r')$ we obtain

$$-r^2\frac{\partial u}{\partial r} + cr = V(r') = r'\frac{\partial v}{\partial r'} - c,$$

or

$$\frac{\partial v}{\partial r'} = cr + cr^2 - r^3\frac{\partial u}{\partial r}.$$

Now

$$v(r') = v(Q) + \int_1^{r'}\frac{\partial v}{\partial r'}\,dr' = v(Q) - c\log r - c(r-1) + \int_1^r r\frac{\partial u}{\partial r}\,dr.$$

Just as we obtained (4), we obtain in this case

$$\Delta^2 v(r') = \frac{1}{r'^2}\left\{\Delta_2[v(Q) - u(Q)] - \left(\frac{\partial u}{\partial r}\right)_Q + c\right\}.$$

The function $w(Q) = v(Q) - u(Q)$ must then satisfy, everywhere on S,

$$\Delta_2 w(Q) = \left(\frac{\partial u}{\partial r}\right)_Q - c.$$

This last equality is impossible, for

$$\iint_S \left\{\left(\frac{\partial u}{\partial r}\right)_Q - c\right\} dS_Q = -4\pi c,$$

while from (5), for $G \equiv 1$ and $H = w$, we have

$$\iint_S \Delta_2 w(Q) dS_Q = 0.$$

THEOREM 3. *If $(\partial u/\partial r) = c \neq 0$ on the open set Ω of the sphere S, and if Σ is a domain containing S and its interior, it is in no case possible to extend u harmonically across Ω into the portion of Σ exterior to S.*

THE UNIVERSITY OF CALIFORNIA

ON THE CLASS OF METRICS DEFINING A METRISABLE SPACE*

H. E. VAUGHAN

Suppose we are given a metrisable† space E. Let M be the class of all allowable metrics on E. Let M_b, M_c, M_B, and M_C be, respectively, the classes of metrics in which the space is bounded, complete, totally bounded, and totally complete. The purpose of this note is to obtain systematically all possible theorems which state the equivalence of some topological property of E (such as compactness, or separability) to the existence or non-existence of metrics having some of the above properties. An example is the well known theorem:

In order that E be compact it is necessary and sufficient that it be complete in every allowable metric.

The problem may also be stated as follows: Using the four definitions as principles of classification and noting the inclusions $M_b \supset M_B \supset M_bM_C$ and $M_c \supset M_C \supset M_cM_B$, we may represent M as the sum of seven disjoint sets: (1) $M - M_b - M_c$, (2) $M_b - M_B - M_bM_c$, (3) $M_c - M_C$

* Presented to the Society, December 28, 1937.

† A topological space will be called *metrisable* if it is possible to define its continuity properties by means of a metric. Any metric which serves this purpose will be called *allowable*, and the space in conjunction with such a metric will be called a *metric space*. A metric space will be called *bounded* if there is a finite upper bound to the distance between any pair of its points. It will be called *complete* if every Cauchy sequence converges. It will be called *totally bounded* if it is, for every positive number e, the sum of a finite number of sets of diameter less than e. It will be called *totally complete* if every bounded set is compact. See C. Kuratowski, *Topologie* I, pp. 82, 87, 91, 196.

$-M_bM_c$, (4) $M_bM_c-M_cM_B-M_bM_C$, (5) $M_B-M_cM_B$, (6) $M_C-M_bM_C$, and (7) M_BM_C. These sets generate a Boolean algebra with 2^7 elements. Any relation among these elements (which is sometimes, but not identically, satisfied) corresponds to some topological property of E, and the problem is to determine this correspondence.

Now, any such relation among the elements may be exhibited as a proposition constructed by the logical operations of conjunction, disjunction, and negation from the seven propositions obtained by setting each of the generating sets equal to the null set. Moreover, by the use of DeMorgan's Law, any such proposition may be expressed as a disjunction of conjunctions, the elements of the latter being chosen from the seven original propositions and their negations. The first step, then, is to obtain the topological equivalents of the 2^{14} conjunctions. Fortunately, only forty-five need be considered, the remainder falling into three classes: (a) those which are never satisfied, (b) those which are always satisfied, and (c) those of the form PQ where P is one of the forty-five and implies Q. It will be readily seen that when the topological equivalents of the above forty-five conjunctions are known, and when it is known that every other conjunction falls into one of the above three classes, it will be a rather easy matter to determine the topological equivalent of any given proposition of the type considered, or to determine whether or not such a proposition which is equivalent to a given property exists.

The topological part of the solution is contained in the following seven theorems and their converses. Proofs are given for the first three theorems. The proofs of the four remaining theorems, as well as of the seven converses, are obvious.

THEOREM 1. *If E is either bounded or complete in every allowable metric, then E is compact.*

PROOF. Suppose that E is not compact, and let (a_i) be a divergent sequence of points of E. Let R be a ray (half-line) with vertex 0, and let b_{2n} be the point of R whose distance from 0 is $1/n$, and b_{2n-1} the point of R whose distance from 0 is $n+1$. The subset of E which consists of the points a_i is closed and is homeomorphic to the subset of R which consists of the points b_i. In fact the correspondence in which corresponding points have identical subscripts is a homeomorphism. Hence* there exists an allowable metric for E such that this correspondence is an isometry. But in such a metric, E is neither bounded nor complete, since the set whose elements are the points

* See F. Hausdorff, *Erweiterung einer Homöomorphie*, Fundamenta Mathematicae. vol. 16 (1930), pp. 352–360.

a_{2i-1} is not bounded, while the sequence (a_{2i}) is a Cauchy sequence which does not converge to any point of E.

THEOREM 2. *If E is either totally bounded or complete in every allowable metric in which it is bounded, then E is compact.*

PROOF. Suppose that E is not compact, and let (a_i) be a divergent sequence of points of E. Let R' be a metric space with points b_i, $(i=1, 2, \cdots)$, and a metric defined as follows: The set consisting of the points b_{2n} is isometric to the similarly named point set in the preceding proof. The distance between any two points b_{2n-1} and b_{2m-1} is zero or one according as n is or is not equal to m. The distance between b_{2n} and b_{2m-1}, in either order, is $(2n-1)/2n$. The triangle axiom is easily verified. As in the preceding proof there exists an allowable metric for E such that the correspondence in which the point a_i is associated with the point b_i is an isometry. From this metric a new metric can be obtained, by the usual transformation $x'=x/(1+x)$, in which the space is bounded. But in this metric, E is neither totally bounded nor complete. For the set whose elements are the points a_{2i-1} is not totally bounded, while the sequence (a_{2i}) is a Cauchy sequence which does not converge to any point of E.

THEOREM 3. *If E allows a metric in which it is complete, and if E is either totally complete or bounded in every such metric, then E is compact.*

PROOF. Suppose that E is not compact, and let (a_i) be a divergent sequence of points of E. Choose a metric for E as in the proof of Theorem 1. There exists* a metric in which E is complete such that the distance in it of any two points of E differs by at most one from the distance assigned to the two points by the metric of Theorem 1. But in this (complete) metric, E is neither totally complete nor bounded, for the set whose elements are the points a_{2i} has a diameter at most equal to two but is not compact, while the set whose elements are the points a_{2i-1} is not bounded.

THEOREM 4. *If E allows a metric in which it is bounded and complete, and if E is either totally bounded or totally complete in at least one such metric, then E is compact.*

THEOREM 5. *If E allows a metric in which it is totally bounded, and if E is complete in every such metric, then E is compact.*

THEOREM 6. *If E allows a metric in which it is totally complete, and if E is bounded in at least one such metric, then E is compact.*

* See C. Kuratowski, *Topologie* I, p. 200.

THEOREM 7. *If E does not allow a metric in which it is both totally bounded and totally complete, then E is not compact.*

From these theorems together with the metric characterizations of separable spaces,* of absolute G_δ's,† and of locally compact, separable spaces‡ it is now possible to determine the topological equivalents of the seven propositions mentioned above and of their negations.

$1. \equiv . M - M_b - M_c = 0$ $. \equiv .$ *E is compact.*
$2. \equiv . M_b - M_B - M_b M_c = 0$ $. \equiv .$ *E is compact.*
$3. \equiv . M_c - M_C - M_b M_c = 0$ $. \equiv .$ *E is compact or not an absolute G_δ.*
$4. \equiv . M_b M_c - M_c M_B - M_b M_C = 0$ $. \equiv .$ *E is compact or not an absolute G_δ.*
$5. \equiv . M_B - M_c M_B = 0$ $. \equiv .$ *E is compact or not separable.*
$6. \equiv . M_C - M_b M_C = 0$ $. \equiv .$ *E is compact or not separable or not locally compact.*
$7. \equiv . M_B M_C = 0$ $. \equiv .$ *E is not compact.*

It is seen that $1. \equiv .2. \equiv . -7$ and that $3. \equiv .4$. Consequently, of the thirty-one conjunctions which contain two or more elements, only the eleven which can be built from 1, 3, 5, 6, and their negations need be listed. These are listed below, with their topological equivalents.

$3 \cdot 5. \equiv .$ *E is compact or not an absolute G_δ and is not separable.*
$3 \cdot -1. \equiv .$ *E is not an absolute G_δ.*
$3 \cdot -5. \equiv .$ *E is separable and not an absolute G_δ.*
$5 \cdot -1. \equiv .$ *E is not separable.*
$5 \cdot -3. \equiv .$ *E is an absolute G_δ and is not separable.*
$6 \cdot -1. \equiv .$ *E is not separable or not locally compact.*
$6 \cdot -3. \equiv .$ *E is an absolute G_δ and is either not separable or not locally compact.*
$6 \cdot -5. \equiv .$ *E is separable and not locally compact.*
$-3 \cdot -5. \equiv .$ *E is separable and an absolute G_δ and is not compact.*
$3 \cdot 5 \cdot -1. \equiv .$ *E is not an absolute G_δ and is not separable.*
$6 \cdot -3 \cdot -5. \equiv .$ *E is separable and an absolute G_δ and is not locally compact.*

The completion of this list makes it easily possible to translate any proposition of the class considered into more familiar topological

* See Kuratowski, loc. cit., p. 91.

† See Kuratowski, loc. cit., p. 215.

‡ See H. E. Vaughan, *On locally compact metrisable spaces*, this Bulletin, vol. 43 (1937), pp. 532–535.

terms. The inverse problem of determining a proposition equivalent to a given topological condition, or showing that one does not exist, may be solved as follows: It will be seen that every condition which appears on the right side of one of the above equivalences is constructed, by means of the three elementary logical operations, from the four propositions:

A: *E is compact.*
B: *E is locally compact and separable.*
C: *E is separable.*
D: *E is an absolute* G_δ.

Furthermore, these four may themselves be obtained from the conditions listed by means of the logical operations. It follows that the topological conditions which may be stated in terms of the existence or non-existence of metrics of the kinds considered are exactly those which can be obtained from the above four by conjunction, disjunction, and negation. To save space only the fourteen conditions involving A, B, and C, but not D, will be stated. The verification, by enumeration, of the fact that these are all offers no difficulty.

A: *E is compact.*
B: *E is locally compact and separable.*
C: *E is separable.*
$-A$: *E is not compact.*
$-B$: *E is not locally compact or not separable.*
$-C$: *E is not separable.*
$B \cdot -A$: *E is locally compact and separable and not compact.*
$C \cdot -A$: *E is separable and not compact.*
$C \cdot -B$: *E is separable and not locally compact.*
$A + -B$: *E is compact or not locally compact or not separable.*
$A + -C$: *E is compact or not separable.*
$A + C \cdot -B$: *E is compact or separable and not locally compact.*
$B + -C$: *E is locally compact or not separable.*
$-C + B \cdot -A$: *E is locally compact and not compact or not separable.*

The proposed problem may now be considered as solved since any equivalence of the kind described in the first paragraph can now be constructed easily, starting with either the metric or the topological condition. It is now also possible to make statements like the following: There is no such equivalence in which the topological condition is local compactness. For the condition "E is locally compact" does not appear in the preceding list.

UNIVERSITY OF ILLINOIS

FORMAL SYNTHESIS OF TWO PERIODIC CORRESPONDENCES, OF PERIOD FIVE AND SEVEN, RESPECTIVELY*

H. S. WHITE

Since the illumination of binary doubly quadratic forms by the theory of elliptic functions, it has been generally the custom to study them in no other light. Even Poncelet's picture of two circles, or a pencil of circles, receives scant attention from the general student. This situation resembles a monopoly and is adverse to the progress of diversified industries. A (2, 2) correspondence between variables (x) and (y) creates sequences of values, each of which is considered first as a y, then as an x. I shall consider one such relation, one which is symmetric and periodic. Its period shall be five.

Denote by $\Phi=0$, or $\Phi(x, y)=\Phi(y, x)=0$, a (2, 2) correspondence, symmetric in x and y. For convenience, we may regard it sometimes as homogeneous in x_1 and x_2, sometimes in y_1 and y_2, or at will as non-homogeneous.

Being symmetric, it has a unique representation quadratic in the two combinations $(x+y)$ and xy, or homogeneous in the three, x_1y_1, $(x_1y_2+x_2y_1)$, x_2y_2. If these combinations be taken as line-coordinates while x, y are point parameters on a conic, the corresponding equation $\Phi=0$ becomes the line-equation of a second conic, and we have Poncelet's picture.

Let five quantities, p, q, r, s, t, constitute a closed series or cycle, each one with its successor forming a pair x, y which satisfies the equation $\Phi=0$. As that symmetric equation has only five constants, it is completely determined by these five pairs. Hence we obtain

$$(1)\qquad F(x, y)=\begin{vmatrix} p^2q^2 & pq(p+q) & (p+q)^2 & pq & p+q & 1 \\ q^2r^2 & qr(q+r) & \cdot & \cdot & \cdot & 1 \\ r^2s^2 & \cdot & \cdot & \cdot & \cdot & 1 \\ s^2t^2 & \cdot & \cdot & \cdot & \cdot & 1 \\ t^2p^2 & \cdot & \cdot & \cdot & \cdot & 1 \\ x^2y^2 & xy(x+y) & (x+y)^2 & xy & x+y & 1 \end{vmatrix}=0.$$

This equation includes extraneous factors, $p-r$, $q-s$, $r-t$, $s-p$, $t-q$. To reduce it by exclusion of these factors and express $\Phi(x, y)$

* Presented to the Society, September 6, 1923, under the title, *Note on five points and a cyclic correspondence.*

$= F(x, y)/\pi(p-r)$, it is most expeditious to use the two conics already mentioned

$$X_1 : X_2 : X_3 = x_1^2 : x_1x_2 : x_2^2,$$
$$U_1 : U_2 : U_3 = x_2y_2 : -(x_1y_2 + x_2y_1) : x_1y_1 \tag{2}$$

and the pascalian equation for the line-conic. This is then reduced by replacing each 3-rowed determinant by three of two rows each. For example,

$$\begin{vmatrix} p^2 & p & 1 \\ q^2 & q & 1 \\ r^2 & r & 1 \end{vmatrix} \equiv -(p-q)(q-r)(r-p). \tag{3}$$

Then the binary identities allow reduction of $F(x, y)$ to any desired standard form such as the following:

$$\begin{aligned} 0 = {} & 2\cdot(pr)(ps)(rs)\cdot(qx)(qy)\cdot(tx)(ty) \\ & + (qr)(st)\cdot(px)(py)\left\{\begin{matrix} (qs)[(rx)(ty) + (ry)(tx)] \\ + (rt)[(qx)(sy) + (qy)(sx)] \end{matrix}\right\}. \end{aligned} \tag{4}$$

One might begin with this standard form of equation and verify by inspection the fact that it is satisfied by each of the five pairs of values (x, y) prescribed. From this explicit equation it can be seen as readily as from the geometric picture that the following is true:

If p, q, r, s, t are real parameters arranged in order of magnitude, then, for example, the passage of x from p to q proceeds pari passu with that of y from q to r.

This simplifies the argument, but eventually the order of arrangement can be arbitrary without impairment of the conclusions.

The correspondence (4) contains at least one cyclic set of five. How is it to be seen that its chains or sequences are all cyclic of period five? For this purpose we attempt to deviate from the initial values p, q, r, s, t, which determine the correspondence, to a neighboring set $p+p', q+q', \cdots, t+t'$, which shall determine the same correspondence, by the *same formal equation*. Since (4) can be regarded as the equation of a line-conic, it is only necessary to make the derived conic intersect it in five points, when the two will necessarily coincide completely. But the pairs $(p, q), (q, r), \cdots$ give five convenient points for comparison. Take the form (1), and vary $p, q, \cdots$ so little that none of the extraneous factors shall vanish. The increment of first order is to yield approximately the equation in (x, y) of our second line-conic. It is

$$0 = p' \frac{\partial F}{\partial p} + q' \frac{\partial F}{\partial q} + \cdots + t' \frac{\partial F}{\partial t} \cdot \tag{5}$$

Examine now $\partial F/\partial p$ for values $x=q$, $y=r$. It vanishes on account of the determinant structure; so also for values (r, s) and for values (s, t), but not for (t, p) nor for (p, q). As further, the increment of the extraneous factor $\prod(p-r)$ (cyclic) is not zero, we have also for the essential factor $\Phi(x, y)$:

The derivative $\partial\Phi(x, y)/\partial p$ becomes zero for three pairs of values (x, y), namely (q, r), (r, s), (s, t); and each of the other derivatives in cyclic order gives in line-coordinates xy, $(x+y)$, 1 the equation of a conic through three of the five points of the given fundamental cycle.

There are therefore five two-term linear equations in $p', q', \cdots, t'$ to be satisfied, namely,

$$\begin{aligned} p' \frac{\partial\phi(p, q)}{\partial p} + q' \frac{\partial\phi(p, q)}{\partial q} &= 0, \\ q' \frac{\partial\phi(q, r)}{\partial q} + r' \frac{\partial\phi(q, r)}{\partial r} &= 0, \\ \cdots\cdots\cdots\cdots\cdots & \end{aligned} \tag{6}$$

In these, the notation intended is

$$\frac{\partial\phi(p, q)}{\partial p} = \left[\frac{\partial}{\partial p}\Phi(x, y)\right]_{x=p, y=q} \cdot$$

Is the determinant of these equations identically zero?

Recur to formula (4); differentiate, and reduce. Our equations become the following:

$$-(p'p)(pq)(pr)(qs)\cdot(qr)(st)(tq) + (q'q)(qp)(tq)(ps)\cdot(pr)(rs)(tp) = 0,$$

or

$$(p'p)\cdot(qr)(qs)(st) + (q'q)(rs)(ps)(tp) = 0,$$

and the four derived by cyclic permutation. The determinant of these five linear equations for (pp'), (qq'), $\cdots$ has two terms only, and they are identical but of opposite sign.

Hence the equations are consistent and determine the ratios of the five differentials $(p'p)$, $(q'q)$, $(r'r)$, $(s's)$, $(t't)$, uniquely.

In this way we see that there is an infinitesimal variation in the

quintette p, q, r, s, t which leaves invariant the equation $\Phi(x, y) = 0$ of the (2, 2) correspondence. As each consecutive set is uniquely determined by the variation of one point, the totality of such quintettes is a linear involution. This is a particular instance of the classical theorem (for example, that of Hurwitz) on closed sets in a (2, 2) correspondence; also of Coble's more general theorem on closed sets in any binary (m, n) correspondence.

The formula of the foregoing problem gives a model which serves, with slight modification, in a problem essentially more intricate. Coble has pointed out that since there exists a (3, 3) correspondence with a closed cycle or period of twice seven elements, there must exist also a (6, 6) correspondence with a closed cycle of seven elements, wherein accordingly each corresponds to all the rest. To construct its explicit expression, let the seven elements be $\alpha, \beta, \gamma, \delta, \epsilon, \zeta, \theta$. Arrange them cyclically in seven triads, as a Δ_7: $\alpha\beta\delta, \beta\gamma\epsilon, \cdots, \theta\alpha\gamma$. Form binary cubics in (u) such as $(u\alpha)(u\beta)(u\gamma)$, whose coefficients shall be taken as quaternary coordinates, $\xi_1' : \xi_2' : \xi_3' : \xi_4'$. Quadric products of these give ten column heads, and the remaining six triads fill the array to seven rows. Fill the last three rows similarly from triads $\pi xy, \rho xy, \sigma xy$, and close the determinant, then divide by the obvious factor $(\pi\rho)(\pi\sigma)(\rho\sigma)$. Call the quotient $D_{10}/(\pi\rho)(\pi\sigma)(\rho\sigma) = \Phi(x, y)$.

The property of this double sextic Φ is that it vanishes for every pair of values $x = \alpha$, $y = \beta$, $\cdots$ and is symmetric in x and y.

The latter is visibly true. As to the former, consider D_{10}. It contains α, β in each of the last three rows; and, by the definition of the triad system, once in some earlier row. Thus since the four quantics $(u\pi)^2(u\alpha)^2(u\beta)^2$, $(u\rho)^2(u\alpha)^2(u\beta)^2$, $(u\sigma)^2(u\alpha)^2(u\beta)^2$, and $(u\gamma)^2(u\alpha)^2(u\beta)^2$ are linearly dependent, so also are the quadrics whose coefficients form four rows in D_{10}. Therefore D_{10} vanishes, and the $\Phi(\alpha, \beta) = 0$.

The equation $\Phi(x, y) = 0$ is satisfied whenever a plane through the two points x, y of a gauche cubic curve (on which seven points represent the parameters $\alpha, \beta, \gamma, \cdots$) osculates the second gauche cubic determined by six planes of the Δ_7.

For such a plane through x, y will have the square of its equation linearly compounded from those given by the last three rows of D_{10}; and, on the other hand, as it osculates the second cubic envelope, it must be linearly compounded from the first seven. Hence there is a linear identity among the ten quadrics corresponding to the rows of the D_{10}; whence follows the above proposition.

VASSAR COLLEGE

DIVISIBILITY OF GENERALIZED FACTORIALS*

BENJAMIN ROSENBAUM

1. **Introduction.** Two different types of expression were obtained by A. M. Legendre† for H, the index of the highest power of the prime p dividing $n!$:

$$H = \left[\frac{n}{p}\right] + \left[\frac{n}{p^2}\right] + \left[\frac{n}{p^3}\right] + \cdots, \tag{1}$$

$$H = \frac{n - s}{p - 1}, \tag{2}$$

where $[a/b]$ denotes the largest integer less than or equal to a/b, and s is the sum of the digits of n to the base p. R. D. Carmichael‡ considered the more general problem of determining H for $\prod_{x=0}^{n-1}(xa+c)$, where a and c are relatively prime positive integers and $a \not\equiv 0 \pmod{p}$. He obtained expressions of type (1) and upper and lower bounds for H. In the present paper a correction is made in the upper bound, new expressions for H of types (1) and (2) are derived, and the results are extended to products where a and c are any positive integers.

2. **Discussion of previous results.** Carmichael used the following method: Set $c_0 = c$, and let i_r be the smallest value of $x \geqq 0$ such that $xa + c_{r-1} \equiv 0 \pmod{p}$, the quotient being c_r. Then $i_r \leqq p-1$. Let $e_0 = n-1$, $e_r = [(e_{r-1} - i_r)/p]$, $(r > 0)$. If $\prod_{x=0}^{n-1}(xa+c_0)$ is divisible by p, it has e_1+1 factors of the form $(mp+i_1)a+c_0$, $(0 \leqq m \leqq [(e_0-i_1)/p])$, each divisible by p. The product of the quotients is $\prod_{x=0}^{e_1}(xa+c_1)$. If this product is divisible by p, it has e_2+1 factors of the form $(mp+i_2)a+c_1$, $(0 \leqq m \leqq [(e_1-i_2)/p])$, each divisible by p. Hence e_2+1 factors of $\prod_{x=0}^{n-1}(xa+c_0)$ are divisible by p^2. If the product of the quotients $\prod_{x=0}^{e_2}(xa+c_2)$ is divisible by p, e_3+1 factors of $\prod_{x=0}^{n-1}(xa+c_0)$ are divisible by p^3. Continue in this manner until a product $\prod_{x=0}^{e_t}(xa+c_t)$ is obtained which is not divisible by p. Then e_t+1 factors of the original product are divisible by p^t and no factors by p^{t+1}. Hence

$$H = \sum_{r=1}^{t}(e_r + 1). \tag{3}$$

* Presented to the Society, April 10, 1936. By a generalized factorial we mean a product of integers forming an arithmetic progression.

† *Théorie des Nombres*, 2d edition, 1808, p. 8.

‡ This Bulletin, vol. 15 (1908–1909), pp. 217–221.

For certain values of a, c_0, and p, one has $c_0=c_1=\cdots=c$ and $i_1=i_2=\cdots=i$. In that case

$$H=\left[\frac{n-1-i+p}{p}\right]+\left[\frac{n-1-i-ip+p^2}{p^2}\right]+\left[\frac{n-1-i-ip-ip^2+p^3}{p^3}\right]+\cdots.$$

In the case of $1\cdot 3\cdot 5\cdot\cdots\cdot(2n-1)$, $i=(p-1)/2$ for $p\neq 2$ and

$$H=\left[\frac{2n-1+p}{2p}\right]+\left[\frac{2n-1+p^2}{2p^2}\right]+\left[\frac{2n-1+p^3}{2p^3}\right]+\cdots. \tag{4}$$

Carmichael also obtained the expression

$$\frac{n-s}{p-1}\leqq H\leqq h+\frac{n-s}{p-1}$$

when n is not a power of p, and $H=(n-1)/(p-1)$ when n is a power of p, where s is the sum of the digits of n to the base p and h is the index of the highest power of $p\leqq n$. The following examples show that these expressions are incorrect: When $a=5$, $c_0=6$, $n=3$, and $p=2$, one has $H=5$ while $h+(n-s)/(p-1)=2$. When $a=2$, $c_0=21$, $n=4$, and $p=3$, one has $H=4$ while $h+(n-s)/(p-1)=2$. When $a=5$, $c_0=1$, $n=4$, and $p=2$, one has $H=5$ while $(n-1)/(p-1)=3$. It will be shown in §8 that the error in the first expression lies in the term h. The second expression was derived from a source containing a similar error. The use of (12) in the above examples gives upper bounds for H of 5, 4, and 5, respectively.

I. Schur* obtained a result equivalent to (4) by the use of a different method. He found $H=\sum_{r=1}^{\infty}[n/p^r+1/2]$.

E. Stridsberg,† considering the same problem as Carmichael, obtained very complicated expressions for H.

3. **Some relations between the letters** c. We shall make use of the following theorem and corollaries:

* Sitzungsberichte der Preussischen Akademie der Wissenschaften, Physikalisch-Mathematische Klasse, 1929, p. 372.

† Arkiv för Matematik, Astronomi och Fysik, vol. 6 (1911), no. 34; summary in Dickson, *History of the Theory of Numbers*, vol. 1, p. 264.

THEOREM. *If c_r and c_s are any two of the letters c, with $s>r$, then c_s is the least integer satisfying the conditions*: (1) $c_s p^{s-r} \equiv c_r \pmod{a}$, (2) $c_s p^{s-r} \geqq c_r$.

PROOF. The theorem is true for c_{r+1}, since i_{r+1} is the least non-negative integer such that $i_{r+1}a + c_r \equiv 0 \pmod{p}$, the quotient being c_{r+1}. Proceed by induction, assuming that c_v is the least integer such that $c_v p^{v-r} \geqq c_r$ and $c_v p^{v-r} \equiv c_r \pmod{a}$. Now $i_{v+1}a + c_v = c_{v+1}p$. Hence c_{v+1} is the least integer such that $c_{v+1}p^{v+1-r} \geqq c_v p^{v-r}$ and $c_{v+1}p^{v+1-r} \equiv c_v p^{v-r} \pmod{a}$. It follows from the properties of c_v that c_{v+1} is the least integer such that $c_{v+1}p^{v+1-r} \geqq c_r$ and $c_{v+1}p^{v+1-r} \equiv c_r \pmod{a}$. The theorem is therefore true for c_{v+1} and consequently for c_s.

COROLLARY 1. *If ϵ is the least positive integer such that $p^\epsilon \equiv 1 \pmod{a}$ and $s>r$, then $c_s = ma + $ residue of $c_r p^{k\epsilon+r-s} \pmod{a}$, where k is any integer such that $k\epsilon + r - s \geqq 0$ and m is the least non-negative integer such that $ma + $ residue $c_r p^{k\epsilon+r-s} \geqq c_r p^{r-s}$. When $c_r < a$, $m = 0$.*

PROOF. The first part of the corollary follows from the theorem, which may be restated in the form: c_s is the least integer greater than or equal to $c_r p^{r-s}$ and congruent to $c_r p^{k\epsilon+r-s}$ modulo a.

To prove the second part of the corollary we make use of the congruence $xp^{s-r} \equiv c_r \pmod{a}$, which has a unique solution $0 \leqq x_1 < a$. When $c_r < a$, $x_1 p^{s-r} \geqq c_r$, otherwise the positive integer $c_r - x_1 p^{s-r}$ is less than a and is congruent to zero modulo a. By the theorem, $x_1 = c_s$. Therefore $c_s < a$ and $m = 0$.

When p is large, the above corollary gives a method for calculating the letters c which is more rapid than that based on the initial determination of i_s as the least non-negative integer such that $i_s a + c_{s-1} \equiv 0 \pmod{p}$. This is especially true when $c_0 < a$.

EXAMPLE. When $c_0 = 29$, $a = 7$, and $p = 11$, $\epsilon = 3$. Then $c_1 = 7m + \text{residue } (29)(11)^{3+0-1} \pmod{7} = 7m + \text{residue } (1)(4)^2 = 7m + 2 = 9$, $(2 < c_0 p^{-1} = 29/11 < 9)$, and $c_2 = 7m + \text{residue } (9)(11)^2 = 7m + 4 = 4$.

COROLLARY 2. *Necessary and sufficient conditions that $c_r = c_s$ are* (1) $c_r \leqq a$, (2) $p^{s-r} \equiv 1 \pmod{a}$.

PROOF. Since c_s is the least integer satisfying the conditions of the theorem, $c_s p^{s-r} = c_r + ja$, where $j \leqq p^{s-r} - 1$. If $c_r > a$, then $c_s p^{s-r} < c_r + c_r(p^{s-r} - 1) = c_r p^{s-r}$, and $c_s < c_r$. Since $c_s p^{s-r} \equiv c_r \pmod{a}$ and c_0 is relatively prime to a, so are all the letters c. Therefore when $c_r = c_s$, we have $p^{s-r} \equiv 1 \pmod{a}$, and the conditions are necessary.

By Corollary 1, when $c_r < a$, $c_s = \text{residue } c_r p^{k\epsilon+r-s} \pmod{a}$. If, in addition, $p^{s-r} \equiv 1 \pmod{a}$, then $c_s = \text{residue } c_r \pmod{a} = c_r$. When $c_r = a$, we have $a = 1$ and $c_s = ma = 1$. Hence the conditions are sufficient.

4. **Expression for H involving the letters i.** Since $\prod_{x=0}^{e_t}(xa+c_t)\not\equiv 0 \pmod{p}$, and $i_{t+1}a+c_t=c_{t+1}p$, it follows that $i_{t+1}>e_t$. Also $i_{t+1}\leqq p-1$. Hence $-1<(e_t-i_{t+1})/p<0$ and

$$e_{t+1}=\left[\frac{e_t-i_{t+1}}{p}\right]=-1.$$

By induction, when $r>t$,

$$e_r=\left[\frac{e_{r-1}-i_r}{p}\right]=-1.$$

Thus (3) is equivalent to

$$H=\sum_{r=1}^{\infty}(e_r+1). \tag{5}$$

Using the values of e_r in §2, substituting that of e_0 in e_1, the resulting value of e_1 in e_2, $\cdots$, we obtain from (5)

$$H=\left[\frac{n-1-i_1+p}{p}\right]+\left[\frac{n-1-i_1-i_2p+p^2}{p^2}\right]+\left[\frac{n-1-i_1-i_2p-i_3p^2+p^3}{p^3}\right]+\cdots. \tag{6}$$

5. **Expression for H involving the letters c.** Consider $i_ra+c_{r-1}=c_rp$. Solving for i_r and substituting in (6) we obtain

$$H=\left[\frac{l}{ap}+\frac{a-c_1}{a}\right]+\left[\frac{l}{ap^2}+\frac{a-c_2}{a}\right]+\left[\frac{l}{ap^3}+\frac{a-c_3}{a}\right]+\cdots, \tag{7}$$

where $l=a(n-1)+c_0$ is the last factor of the product $\prod_{x=0}^{n-1}(xa+c_0)$.

Since $e_r+1\geqq 1$ for $r\leqq t$ and $e_r+1=0$ for $r>t$, all terms of (5), (6), and (7) are zero after the first zero term.

When $a=1$ or 2 and $a\not\equiv 0 \pmod{p}$, we have $p\equiv 1 \pmod{a}$. By Corollary 2, when $c_0\leqq a$, $c_0=c_1=\cdots=c$ and (7) give (1) or (4).

When $a=3$, 4, or 6 and $a\not\equiv 0 \pmod{p}$, we have $p\equiv 1$ or $p\equiv -1 \pmod{a}$. When $c_0<a$ and $p\equiv 1$, $c_0=c_1=\cdots=c$. When $c_0<a$ and $p\equiv -1$, since $p^2\equiv 1 \pmod{a}$, $c_0=c_2=c_4=\cdots$. By Corollary 1, $c_1=$residue of $c_0p \pmod{a}$. Hence $c_1\equiv -c_0\equiv a-c_0 \pmod{a}$, and $c_1=a-c_0=c_3=c_5=\cdots$.

6. **Expression for H involving digits of n to base p.** Let $n=d_kp^k+d_{k-1}p^{k-1}+\cdots+d_1p+d_0$, and let $s=d_0+d_1+\cdots+d_k$, with $0\leqq d_r\leqq p-1$. On substituting the above value of n in (6) we obtain

$$H=\sum_{r=1}^{\infty}\left[\frac{d_kp^k+d_{k-1}p^{k-1}+\cdots+d_rp^r}{p^r}+\frac{p^r+d_{r-1}p^{r-1}+\cdots+d_1p+d_0-i_rp^{r-1}-\cdots-i_2p-i_1-1}{p^r}\right].$$

We shall designate the second term in the brackets by F_r. When $d_{r-1}p^{r-1}+d_{r-2}p^{r-2}+\cdots+d_0\geqq i_rp^{r-1}+i_{r-1}p^{r-2}+\cdots+i_1+1$, we obtain $1\leqq F_r<2$. Since each d and each i is less than or equal to $p-1$, this will occur when and only when $d_{r-1}>i_r$, or

$$d_{r-1}=i_r \quad \text{and} \quad d_{r-1-b}>i_{r-b}, \tag{8}$$

where $r-1-b\geqq 0$ and d_{r-1-b} is the first d of lower subscript than d_{r-1} which is not equal to the corresponding i. (The letter i_r corresponds to d_{r-1}. Though $d_{k+u}=0$ when $u\geqq 1$, it is possible to have the corresponding letter $i=0$ and $F_{k+v}\geqq 1$, $v\geqq 1$.) When $d_{r-1}p^{r-1}+d_{r-2}p^{r-2}+\cdots+d_0<i_rp^{r-1}+i_{r-1}p^{r-2}+\cdots+i_1+1$, we have $0\leqq F_r<1$. From the above it follows that $H=\sum_{r=1}^{\infty}[n/p^r]+\sum_{r=1}^{\infty}[F_r]$ and finally that

$$H=\frac{n-s}{p-1}+g, \tag{9}$$

where g is the number of values of $r\geqq 1$ for which $d_{r-1}\geqq i_r$, the equality sign being used only when the conditions of (8) are fulfilled.

In the case of $n!$, $i=p-1$. Hence $g=0$ and $H=(n-s)/(p-1)$.

In the case of $1\cdot 3\cdot 5\cdot\cdots\cdot(2n-1)$, $i=(p-1)/2$ and g is the number of values of $r\geqq 0$ for which $d_r\geqq(p-1)/2$, with the restriction on the equality sign.

EXAMPLE. This example illustrates the use of (9). Consider the product (22)(27)(32)(37)(42) with $p=3$. From $i_ra+c_{r-1}=c_rp$ we obtain $i_1=1$, $i_2=0$, $i_3=0$, $i_4=1$; and $n=5=(1)(3)+(2)$. Hence $d_0=2$, $d_1=1$; $d_r=0$, $r>1$. Since $d_0>i_1$, $d_1>i_2$, $d_2=i_3$, and $d_4<i_4$, we have $g=3$. $H=(5-3)/2+3=4$.

7. **Expression for H involving digits of $l=a(n-1)+c_0$ to base p.** Let $l=\delta_kp^k+\delta_{k-1}p^{k-1}+\cdots+\delta_0$ and $\sigma=\delta_0+\delta_1+\cdots+\delta_k$, with $0\leqq\delta_r\leqq p-1$. Since $l\leqq p^{k+1}-1$ and $c_r\geqq 1$, all terms of (7) beyond $[l/ap^k+(a-c_k)/a]$ are zero. Hence

$$H = \sum_{r=1}^{\lambda} \left[\frac{a(n-1) + c_0 + p^r(a - c_r)}{ap^r} \right]$$

$$= \sum_{r=1}^{\lambda} \left[\frac{N_r}{ap^r} + \frac{D_{r-1} + ap^r - R_{r-1}p^r}{ap^r} \right],$$

where $D_{r-1} = \delta_{r-1}p^{r-1} + \delta_{r-2}p^{r-2} + \cdots + \delta_0$. Here R_{r-1} is the residue ($\geqq 1$ and $\leqq a$) of $p^{k\epsilon - r}D_{r-1} \pmod a$, ϵ is the least positive exponent such that $p^\epsilon \equiv 1 \pmod a$, k is an integer such that $k\epsilon - r \geqq 0$, and $N_r = a(n-1) + c_0 - c_r p^r - D_{r-1} + R_{r-1}p^r$. By observing that $a(n-1) + c_0 - D_{r-1} = \delta_\lambda p^\lambda + \cdots + \delta_r p^r$, $c_r p^r - c_0 \equiv 0 \pmod a$ (see the theorem of §3), and $R_{r-1}p^r - D_{r-1} \equiv p^{k\epsilon - r}D_{r-1}p^r - D_{r-1} \equiv 0 \pmod a$, we see that $N_r \equiv 0 \pmod{ap^r}$.

Also because $D_{r-1} \leqq p^r - 1$ and $1 \leqq R_{r-1} \leqq a$, we see that

$$0 \leqq \frac{D_{r-1} + ap^r - R_{r-1}p^r}{ap^r} < 1.$$

Therefore

$$H = \sum_{r=1}^{\lambda} \frac{N_r}{ap^r}$$

$$= \sum_{r=1}^{\lambda} \left(\frac{\delta_\lambda p^{\lambda - r} + \delta_{\lambda-1}p^{\lambda-1-r} + \cdots + \delta_{r+1}p + \delta_r}{a} + \frac{R_{r-1} - c_r}{a} \right)$$

$$= \sum_{r=1}^{\lambda} \left(\frac{\delta_r(p^{r-1} + p^{r-2} + \cdots + 1)}{a} + \frac{R_{r-1} - c_r}{a} \right)$$

$$= \sum_{r=1}^{\lambda} \left(\frac{\delta_r(p^r - 1)}{a(p-1)} + \frac{R_{r-1} - c_r}{a} \right),$$

and finally

$$H = \frac{l - \sigma}{a(p-1)} + \sum_{r=1}^{\lambda} \frac{R_{r-1} - c_r}{a}. \tag{10}$$

In the case of $n!$, we have $a = 1$, $c = 1$, $\epsilon = 1$, $R_{r-1} = 1$, and $\sum_{r=1}^{\lambda}(R_{r-1} - c_r)/a = 0$. Therefore $H = (n-s)/(p-1)$.

In the case of $1 \cdot 3 \cdot 5 \cdot \cdots \cdot (2n-1)$, we have $a = 2$, $c = 1$, and $\epsilon = 1$. Then $R_{r-1} = 1$ when D_{r-1} is odd; $R_{r-1} = 2$ when D_{r-1} is even. Hence $H = (2n - \sigma - 1)/2(p-1) + e/2$, where e is the number of values of r, $(1 \leqq r \leqq \lambda)$, for which D_{r-1} is even. When $l = p^\lambda$, $\sigma = 1$ and $e = \lambda$. Therefore $H = (n-1)/(p-1) + \lambda/2$.

EXAMPLE. This example illustrates the use of (10). Determine H

for (22)(27)(32)(37)(42) with $p=3$. We obtain $\epsilon=4$, $l=42=(1)(3)^3+(1)(3)^2+(2)(3)+0$; $D_0=0$, $D_1=6$, $D_2=15$; $R_0=5$, $R_1=$residue $(3)^{4-2}(6) \pmod 5 = 4$, $R_2=5$. From $i_r a+c_{r-1}=c_r p$, we obtain $c_1=9$, $c_2=3$, and $c_3=1$. Then $H=(42-4)/(5)(2)+(14-13)/5=4$.

8. **Upper and lower bounds of H.** The terms of (5) and (6) vanish after the tth term, where t has the same meaning as in (3). We have $0 \leqq i_r \leqq p-1$. Substituting the limiting values of i_r in (6) we obtain

$$\left[\frac{n}{p}\right]+\left[\frac{n}{p^2}\right]+\cdots \leqq H \leqq \left[\frac{n-1}{p}\right]+\left[\frac{n-1}{p^2}\right]+\cdots+t. \tag{11}$$

It is evident from §2 that t is the index of the highest power of p dividing any one factor of $\prod_{x=0}^{n-1}(xa+c_0)$. Hence $t \leqq \lambda$, the index of the highest power of $p \leqq l=a(n-1)+c_0$. However t may exceed h, the index of the highest power of $p \leqq n$. If α is the index of the highest power of p exactly dividing n, and β is any integer $\geqq 0$, then $[n/p^\beta]=[(n-1)/p^\beta]+1$ for $\beta \leqq \alpha$, and $[n/p^\beta]=[(n-1)/p^\beta]$ for $\beta>\alpha$. Substituting these results in (11), we have

$$\left[\frac{n}{p}\right]+\left[\frac{n}{p^2}\right]+\cdots \leqq H \leqq \left[\frac{n}{p}\right]+\left[\frac{n}{p^2}\right]+\cdots+\lambda-\alpha,$$

or

$$\frac{n-s}{p-1} \leqq H \leqq \frac{n-s}{p-1}+\lambda-\alpha. \tag{12}$$

9. **Values of H when a and c_0 are any positive integers.** If a and c_0 are not relatively prime let d be their greatest common divisor, with $a=a'd$ and $c_0=c'd$. Then $\prod_{x=0}^{n-1}(xa+c_0)=d^n\prod_{x=0}^{n-1}(xa'+c')$. If H, H', and h_d are the indices of the highest powers of p dividing $\prod_{x=0}^{n-1}(xa+c_0)$, $\prod_{x=0}^{n-1}(xa'+c')$, and d, respectively, then $H=H'+nh_d$.

When a and c_0 are relatively prime and $a \equiv 0 \pmod p$, $xa+c_0$ is not divisible by p and $H=0$.

YALE UNIVERSITY

REMARKS ON THE CLASSICAL INVERSION FORMULA FOR THE LAPLACE INTEGRAL

D. V. WIDDER AND N. WIENER

If a function $f(s)=f(\sigma+i\tau)$ is defined for $\sigma>0$ by the Laplace integral

$$f(s) = \int_0^\infty e^{-st}\phi(t)dt, \tag{1}$$

then the classical inversion formula is

$$\phi(t) = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} f(s)e^{st}ds, \qquad c > 0, t > 0. \tag{2}$$

Conditions for the validity of this formula have frequently been discussed. However, the authors know of no adequate treatment* of the case when $\phi(t)$ belongs to L^2 in $(0, \infty)$:

$$\int_0^\infty |\phi(t)|^2 dt < \infty. \tag{3}$$

We employ here the usual notation,

$$\underset{a\to\infty}{\text{l.i.m.}}\ \phi_a(t) = \phi(t),$$

to mean that $\phi_a(t)$ and $\phi(t)$ belong to L^2 in $(-\infty, \infty)$ and that

$$\lim_{a\to\infty}\int_{-\infty}^{\infty} |\phi_a(t) - \phi(t)|^2 dt = 0.$$

It is clear first that if (3) holds then (1) converges absolutely for $\sigma>0$, since

$$|f(\sigma + i\tau)|^2 = \left|\int_0^\infty e^{-st}\phi(t)dt\right|^2 \leqq \int_0^\infty e^{-2\sigma t}dt \int_0^\infty |\phi(t)|^2 dt.$$

Moreover, by the Plancherel theorem regarding Fourier transforms,

$$\underset{a\to\infty}{\text{l.i.m.}}\int_0^a e^{-i\tau t}\phi(t)dt$$

* But compare G. Doetsch, *Bedingungen für die Darstellbarkeit einer Funktion als Laplace Integral und eine Umkehrformel für die Laplace-Transformation*, Mathematische Zeitschrift, vol. 42 (1936), p. 272, Theorem 1.

exists. We denote it by $f(i\tau)$. The same theorem gives us at once that

$$\text{l.i.m.}_{a\to\infty} \frac{1}{2\pi}\int_{-a}^{a} f(i\tau)e^{i\tau t}d\tau = \begin{cases}\phi(t), & t>0,\\ 0, & t<0,\end{cases}$$

or

$$\text{(4)}\qquad \text{l.i.m.}_{a\to\infty} \frac{1}{2\pi i}\int_{-ia}^{ia} f(s)e^{st}ds = \begin{cases}\phi(t), & t>0,\\ 0, & t<0.\end{cases}$$

Hence (2) with $c=0$ is valid in the sense of (4). However, if $c>0$, (2) is no longer valid in this sense.

If $c>0$, it is again clear from the Plancherel theorem that

$$\text{l.i.m.}_{a\to\infty} \frac{1}{2\pi}\int_{-a}^{a} f(c+i\tau)e^{i\tau t}d\tau = \begin{cases}e^{-ct}\phi(t), & t>0,\\ 0, & t<0.\end{cases}$$

But this does not imply that

$$\text{l.i.m.}_{a\to\infty} \int_{c-ia}^{c+ia} f(s)e^{st}ds = \begin{cases}\phi(t), & t>0,\\ 0, & t<0,\end{cases}$$

unless $f(s)$ is identically zero. For, set $\phi_a(t)=\int_{-a}^{a} f(c+i\tau)e^{i\tau t}d\tau$. It will be sufficient to show that

$$\text{(5)}\qquad \int_{-\infty}^{\infty} e^{2ct}\,|\,\phi_a(t)\,|^2dt = \infty$$

for some a. Choose a so that

$$\text{(6)}\qquad f(c+ia) \neq f(c-ia).$$

This is possible, for otherwise we should have by use of (1) that

$$\int_0^{\infty} e^{-ct}\phi(t)\sin at\,dt = 0$$

for all a. By the uniqueness theorem for the Fourier sine transform this would imply that $\phi(t)$ is equivalent to zero and that $f(s)$ is identically zero. In fact we see that (6) may be satisfied for some a in every interval however small.

An integration by parts of the integral defining $\phi_a(t)$ gives

$$\phi_a(t) = \frac{f(c+ia)e^{iat}-f(c-ia)e^{-iat}}{it} - \frac{1}{t}\int_{-a}^{a} e^{i\tau t}f'(c+i\tau)d\tau$$

$$= \frac{f(c+ia)e^{iat}-f(c-ia)e^{-iat}}{it} + o\left(\frac{1}{|t|}\right), \qquad |t|\to\infty,$$

$$= \frac{e^{iat}}{it}\left[\{f(c+ia) - f(c-ia)\} + f(c-ia)(1-e^{-2iat})\right] + o\left(\frac{1}{|t|}\right).$$

Hence

$$t^2|\phi_a(t)|^2 \geqq \left[\frac{k}{2} - 2l\,|\sin at|\right]^2, \qquad t > t_0,$$

where t_0 is a sufficiently large positive number and

$$k = |f(c+ia) - f(c-ia)| \neq 0, \qquad l = |f(c-ia)|.$$

Since $2l|\sin at| < k/4$ in an interval of length δ, say, about $t=0$ and in intervals congruent to this one, modulo π/a, it is clear that the integrand of (5) exceeds $k^2e^{2ct}/16t^2$ in infinitely many intervals of length δ. This is sufficient to insure the divergence of the integral.

We collect our results in the following form:

THEOREM. *If $\phi(t)$ belongs to L^2 in $(0, \infty)$, then it has a Laplace transform $f(\sigma+i\tau)$ defined for $\sigma>0$ by the absolutely convergent integral*

$$f(\sigma + i\tau) = \int_0^\infty e^{-(\sigma+i\tau)t}\phi(t)dt,$$

and for $\sigma=0$ by

$$f(i\tau) = \underset{a\to\infty}{\text{l.i.m.}} \int_0^a e^{-i\tau t}\phi(t)dt.$$

The inversion formula

$$\underset{a\to\infty}{\text{l.i.m.}} \frac{1}{2\pi i}\int_{c-ia}^{c+ia} f(s)e^{st}ds = \begin{cases}\phi(t), & t>0,\\ 0, & t<0,\end{cases}$$

is false ($f(s)\not\equiv 0$) for $c>0$ and valid for $c=0$. For all $c\geqq 0$

$$\underset{a\to\infty}{\text{l.i.m.}} \frac{1}{2\pi}\int_{-a}^{a} f(c+i\tau)e^{i\tau t}d\tau = \begin{cases}e^{-ct}\phi(t), & t>0,\\ 0, & t<0.\end{cases}$$

HARVARD UNIVERSITY AND
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

NON-CYCLIC ALGEBRAS WITH PURE MAXIMAL SUBFIELDS*

A. A. ALBERT

One of the most elementary consequences of the assumption that a normal division algebra A is cyclic of degree n over its centrum K is that A contains a quantity j whose minimum equation is $\omega^n = \gamma$ in K. In 1933 I conjectured the truth of the converse proposition. The proof is easily reducible† to the case where n is a power p^e of a prime p. Let q be the characteristic of K. I succeeded in proving the theorem for $q = p$, e arbitrary,‡ as well as for $q \neq p$, $e = 1$.§ There remained the case $q \neq p$, $e \geqq 2$.

My hope for the truth of the theorem was heightened by H. Hasse's remark‖ that it would provide an essential simplification of the arithmetic existence properties required for the proof of the theorem that all normal division algebras over an algebraic number field are cyclic. However this hope is at an end. For the conjecture is actually false in the remaining case. This is shown by a demonstration of the validity of the following theorem:

THEOREM. *Let x, y, z be independent indeterminates over a field F of real numbers,¶ $K = F(x, y, z)$. Then there exist non-cyclic normal division algebras of degree and exponent four over K, each with a subfield $K(j)$ of degree four over K such that $j^4 = \gamma$ in K.*

Our example is obtained from a class of non-cyclic algebras given in my paper in the Transactions of this Society, vol. 35 (1933), pp.

* Presented to the Society, February 26, 1938.

† For, every A is a direct product of division algebras A_i of degrees $n_i = p_i^{e_i}$ for distinct primes p_i, and A is cyclic if and only if the A_i are cyclic. Moreover the field defined by $\omega^n = \gamma$ splits A if and only if the fields defined by $\omega^{n_i} = \gamma$ split the A_i. For references tc the results used see M. Deuring's *Algebren*.

‡ Transactions of this Society, vol. 39 (1936), pp. 183–188.

§ Ibid., vol. 36 (1934), pp. 885–892.

‖ In a letter to the author. The arithmetic existence theorem is that of W. Grünwald, Journal für die reine und angewandte Mathematik, vol. 169 (1933), pp. 103–107. The proof of Hasse applicable for the case $n = p$ is as follows. Assume that γ is in K and is exactly divisible by the first power of P for every prime ideal P of K such that the P-index of A is not unity. Let also γ be negative for all real fields conjugate to K. Then $K(\gamma^{1/n})$ splits A and is equivalent to a maximal subfield of A. When n is a prime p this implies that A is cyclic.

¶ Our existence theorem is for F non-modular. It seems likely that a modification can be made with F of characteristic any $p > 2$.

112–121. We shall require only slight modifications of the choice of parameters of that paper, and shall be so closely concerned with its content that we shall refer to its equations, lemmas, and theorems by the numbers used therein. Any new equations and lemmas will be numbered consecutively with those of this earlier paper, and we shall indicate modified results by an accent.

Put $\epsilon_1=\gamma_1=0$ in (37). This necessitates the deletion of (29) and gives

$$\rho = -\epsilon_5^2 e. \tag{36'}$$

From (2) we have

$$j_1^2 = g_1 = \gamma_2 u, \qquad j_1^4 = \gamma_2^2\rho = \gamma \text{ in } K. \tag{43}$$

Then j_1 is the quantity j of our theorem. We shall assume (23)–(27), (28), (30)–(34), but shall make some further restrictions. These will be necessary in order that the property proved in the original §9 shall again hold.

Observe first that the condition (7) that our algebra A have exponent four has now become

$$\alpha_1^2 - \alpha_2^2\sigma - (\gamma_2\alpha_3\epsilon_5)^2 e = 0 \tag{7'}$$

for $\alpha_1, \alpha_2, \alpha_3$ in K only if $\alpha_1=\alpha_2=\alpha_3=0$. But this is equivalent to (40), and the proof of §8 depends only upon (33), (34). Hence this proof is valid, and A is a normal division algebra of degree and exponent four with a maximal subfield $K(j)$, $(j^4=\gamma \text{ in } K)$.

We prove that A is non-cyclic by showing that $A\times L$ is a division algebra for every quadratic field $L=K(q)$, $(q=\delta_1^2+\delta_2^2)$.* Observe that it is not possible to prove this by showing that $A\times L$ has exponent four over L. For it is known† that every normal division algebra A of degree four over K has exponent two over an existing field L of our type. Our theorem will thus imply the following corollary:

COROLLARY. *There exist normal division algebras A of degree and exponent four over K and quadratic fields L such that L splits A^2 but $A\times L$ is a division algebra.*

Observe now that (9) becomes $Q=\alpha_1^2+\alpha_2^2\rho-\sigma(\alpha_3^2+\alpha_4^2\rho)$. The vanishing of Q for α_i not all zero is equivalent to the non-trivial vanishing of $\alpha_1^2+\alpha_2^2\rho-\alpha_3^2\sigma$, since $\alpha_3^2+\alpha_4^2\rho$ is the norm form of a quadratic field. But this is precisely the condition (7'). Hence the sufficient

* This is as in the earlier paper referred to above.

† A consequence of Theorem 7 of the author's paper in the Transactions of this Society, vol. 34 (1932), pp. 363–372.

condition of (9) and §9 is not satisfied, and we shall have to investigate the question more deeply. Observe finally that the proof in §10 depends only upon the z degrees of our quantities as given in (23)–(27) and is valid provided that we can prove the result of §9, namely, that the algebra $B \times L$ is a division algebra. To do this we first make the additional assumptions

(43) γ_2 of odd y degree, $\epsilon_5\gamma_2$ of odd y degree.

We then prove the following lemma:

LEMMA 11. *There exist polynomials a, b in $F[x, z]$ such that a^2+b^4 is not the square of any quantity of the field $F(x, z)$.*

For as in Lemma 2 the equation $a^2+b^4=c^2$ for c in $F(x, z)$ implies that c is a polynomial in $F[x, z]$. The condition is easily seen to be satisfied by $a=x^2$, $b=z$.

Our choice of γ_2, γ_3, ϵ_5, γ_4, γ_6 in (30), (31), (43) now implies the truth of the following lemma:

LEMMA 12. *The y degree of ρ is a multiple of four, and its y-leading coefficient is $-a^2$ with a an arbitrary polynomial of $F[x, z]$.*

Choose a as in Lemma 11 with a corresponding to b, and put

$$\phi = (by^n)^4 - \rho, \tag{44}$$

where $4n$ is the y degree of ρ. Then ϕ also has y degree $4n$ and y-leading coefficient

$$\phi_y = b^4 + a^2 \neq c^2 \tag{45}$$

for any c of $F[x, z]$. Observe our introduction of the notation α_y for the y-leading coefficient of any quantity α of $F[y, x, z]$.

We next prove the lemma:

LEMMA 13. *Let $\alpha_1, \cdots, \alpha_8$ be variables over K, and let*

$$S = \sigma(\alpha_1^2 - \alpha_2^2\phi) - (\alpha_7^2 + \alpha_8^2), \quad T = \gamma_2[\alpha_3^2 + \alpha_4^2\rho\phi - \sigma(\alpha_5^2 + \alpha_6^2\rho\phi)].$$

Then the form $S-T$ is not a null form over K.

For let $S-T$ be a null form so that $S=T$ for α_i in K, not all zero. There is clearly no loss of generality if we assume the α_i polynomials in x, y, z. The formal y degree of $\alpha_3^2+\alpha_4^2\rho\phi$ is even, and its y-leading coefficient is evidently a sum of three squares. Similarly the y-leading coefficient of $\sigma(\alpha_5^2+\alpha_6^2\rho\phi)$ is σ_y multiplied by a sum of three squares. But by (28) σ_y has odd x degree. Hence the sum of the y-leading coefficients of $\alpha_3^2+\alpha_4^2\rho\phi-\sigma(\alpha_5^2+\alpha_6^2\rho\phi)$ is zero only if $\alpha_3=\alpha_4=\alpha_5=\alpha_6$

$=0$. But (43) then implies that T either has odd y degree or is zero.

The y-leading coefficient of $\alpha_1^2-\alpha_2^2\phi$ is a sum of terms such as $\alpha_{1y}^2-\alpha_{2y}^2(b^4+a^2)$ and cannot be zero by our choice in Lemma 11 unless $\alpha_1=\alpha_2=0$. Hence the y-leading coefficient of $\sigma(\alpha_1^2-\alpha_2^2\rho)$ has odd x degree or is zero, that of $\alpha_3^2+\alpha_4^2$ has even x degree, and S clearly has even y degree. This shows that $S=T$ implies that $T\equiv 0$, $\alpha_3=\alpha_4=\alpha_5=\alpha_6=0$. Our proof also implies that $\alpha_1=\alpha_2=\alpha_7=\alpha_8=0$, a contradiction. This proves our lemma.

We now replace the argument of §9 by the property that $B\times L$ is not a division algebra if and only if

$$\delta_1^2+\delta_2^2=\beta_1^2\sigma+\beta_2^2 g_1-\beta_3^2\sigma q, \tag{46}$$

for β_1, β_2, β_3 in the centrum $K(u)$ of B, $L=K(q)$, $q^2=\delta_1^2+\delta_2^2$, and δ_1, δ_2 in K. This is a well known property of quaternion algebras.* We write $\beta_i=\xi_i+u\eta_i$ and equate coefficients of u, obtaining the equivalent pair of conditions

$$\delta_1^2+\delta_2^2=(\xi_1^2+\eta_1^2\rho)\sigma+2(\xi_2\eta_2-\sigma\xi_3\eta_3)\gamma_2\rho, \tag{47}$$

$$0=2\xi_1\eta_1\sigma+[(\xi_2^2+\eta_2^2\rho)-\sigma(\xi_3^2+\eta_3^2\rho)]\gamma_2, \tag{48}$$

for ξ_i, η_i, δ_1, δ_2, not all zero, in K. Multiply (48) by λ^2 and add to (47). We obtain

$$\begin{aligned}\delta_1^2+\delta_2^2=&[(\xi_1+\lambda^2\eta_1)^2+\eta_1^2(\rho-\lambda^4)]\sigma+\gamma_2[(\lambda\xi_2+\rho\eta_2\lambda^{-1})^2\\&+\eta_2^2\lambda^{-2}(\lambda^4-\rho)\rho-\sigma(\lambda\xi_3+\rho\eta_3\lambda^{-1})^2-\sigma\eta_3^2\lambda^{-2}(\lambda^4-\rho)\rho].\end{aligned} \tag{49}$$

Put $\lambda=by^n$, $\delta_1=\alpha_1$, $\delta_2=\alpha_8$, $\xi_1+\lambda^2\eta_1=\alpha_1$, $\eta_1=\alpha_2$, $\lambda\xi_2+\rho\eta_2\lambda^{-1}=\alpha_3$, $\eta_2\lambda^{-1}=\alpha_4$, $\lambda\xi_3+\rho\lambda_3\lambda^{-1}=\alpha_5$, $\eta_3\lambda^{-1}=\alpha_6$, and obtain the form $S-T$ of (45). By Lemma 13 the form $S-T$ is not a null form. Hence neither is (49); and (47), (48) cannot have a simultaneous solution. This proves our final result.

THE UNIVERSITY OF CHICAGO

* It is a corollary of the theorem stating that if L is a field of degree n over K, and if A is a division algebra of degree n over K, then L splits A if and only if L is equivalent to a subfield of A. For our quaternion algebras these are fields $K(\beta)$ with $\beta=\beta_1 v+\beta_2 j_1+\beta_3 v j_1$, and we must have $\beta^2=\delta_1^2+\delta_2^2$ as in (46).

THE TRANSFORMS OF FUCHSIAN GROUPS

P. K. REES

This paper will deal with Fuchsian linear transformations that have the unit circle

$$Q_0(z) \equiv z\bar{z} - 1 = 0$$

as principal circle. The transformations have the form

$$T(z) = \frac{az + \bar{c}}{cz + \bar{a}}, \qquad a\bar{a} - c\bar{c} = 1. \tag{1}$$

We shall develop several theorems concerning the relative size of the isometric circle of any transformation T of the group and of the isometric circle of

$$\begin{aligned} S(z) &= GTG^{-1}(z) \\ &= \frac{(-\alpha\bar{\alpha}a + \alpha\nu\bar{c} - \bar{\alpha}\bar{\nu}c + \nu\bar{\nu}\bar{a})z + \alpha\bar{\nu}a - \alpha^2\bar{c} + \bar{\nu}^2c - \alpha\bar{\nu}\bar{a}}{(-\bar{\alpha}\nu a + \nu^2\bar{c} - \bar{\alpha}^2c + \bar{\alpha}\nu\bar{a})z + \nu\bar{\nu}a - \alpha\nu\bar{c} + \bar{\alpha}\bar{\nu}c - \alpha\bar{\alpha}\bar{a}} \\ &= \frac{Az + \overline{C}}{Cz + \overline{A}}, \qquad A\overline{A} - C\overline{C} = 1, \end{aligned} \tag{2}$$

in which

$$G(z) = \frac{\alpha z + \bar{\nu}}{\nu z + \bar{\alpha}}, \qquad \alpha\bar{\alpha} - \nu\bar{\nu} = 1,$$

is variable and T fixed.

The concept of the isometric circle of a linear transformation was introduced in 1926 by L. R. Ford. Because of his work with it, the isometric circle now plays a fundamental part in work with groups of linear transformations. By definition* the isometric circle of a linear transformation is the complete locus of points in the neighborhood of which lengths and areas are unaltered in magnitude by the transformation. For convenience we shall sometimes use the designation $Q(z)$ for a line or circle $Q(z) = 0$.

THEOREM 1. *A necessary and sufficient condition that the radii r_s and r_t of the isometric circles of S and T be equal is that the center w of the isometric circle of the transformation G be either on the line $Q_1(z)$ through the center g_t of the isometric circle of T and the center g_t' of the isometric circle of T^{-1} or on the inverse $Q_2(z)$ of $Q_1(z)$ in $Q_0(z)$.*

* L. R. Ford, *Automorphic Functions*, p. 25.

PROOF. The equation of Q_1 is

$$Q_1(z) = cz - \bar{c}\bar{z} - (a - \bar{a}) = 0.$$

The equation of Q_2, which is the inverse of Q_1 in Q_0, is

$$Q_2(z) = cz - \bar{c}\bar{z} - (a - \bar{a})z\bar{z} = 0.$$

If z lies on either Q_1 or Q_2, we have

$$(3)\quad \begin{aligned} Q_1(z)Q_2(z) = &- (a - \bar{a})(cz - \bar{c}\bar{z})(z\bar{z} + 1) + z\bar{z}[(a - \bar{a})^2 - 2c\bar{c}] \\ &+ c^2z^2 + \bar{c}^2\bar{z}^2 = 0. \end{aligned}$$

Now, using the values of r_t and r_s from (1) and (2), we have

$$(4)\quad \frac{1}{r_s^2} - \frac{1}{r_t^2} = C\overline{C} - c\bar{c} = - (a - \bar{a})(- \alpha\nu\bar{c} + \bar{\alpha}\bar{\nu}c)(\alpha\bar{\alpha} + \nu\bar{\nu})$$

$$- \alpha\bar{\alpha}\nu\bar{\nu}(a - \bar{a})^2 - \alpha^2\nu^2\bar{c}^2 - \bar{\alpha}^2\bar{\nu}^2c^2 + (\alpha^2\bar{\alpha}^2 + \nu^2\bar{\nu}^2 - 1)c\bar{c}$$

$$(5)\quad \begin{aligned} &= - (a - \bar{a})(- \alpha\nu\bar{c} + \bar{\alpha}\bar{\nu}c)(\alpha\bar{\alpha} + \nu\bar{\nu}) \\ &\quad - \alpha\bar{\alpha}\nu\bar{\nu}[(a - \bar{a})^2 - 2c\bar{c}] - \alpha^2\nu^2\bar{c}^2 - \bar{\alpha}^2\bar{\nu}^2c^2, \end{aligned}$$

which vanishes if and only if $r_s = r_t$.

Dividing (5) equated to zero by $-(\nu\bar{\nu})^2$, replacing $-\bar{\alpha}/\nu$ by w, and comparing the result with (3), we see that $r_s = r_t$ if and only if w lies on $Q_1(z)$ or $Q_2(z)$.

THEOREM 2a. *A necessary and sufficient condition that $r_s < r_t$ ($r_s > r_t$) is that $z = w$, substituted in the expression for Q_1Q_2, make that expression negative (positive).*

PROOF. $r_s < r_t$ ($r_s > r_t$) according as $1/r_s^2 - 1/r_t^2 > 0$ ($1/r_s^2 - 1/r_t^2 < 0$). However,

$$Q_1Q_2 = (1/r_s^2 - 1/r_t^2)/(-\nu\bar{\nu})^2.$$

Therefore, $Q_1Q_2 < 0$ or $Q_1Q_2 > 0$ according as $1/r_s^2 - 1/r_t^2 > 0$ or $1/r_s^2 - 1/r_t^2 < 0$.

THEOREM 2b. *A necessary and sufficient condition that $r_s > r_t$ is that w be outside Q_2 and on the opposite side of Q_1 from the origin; a necessary and sufficient condition that $r_s < r_t$ is that w be inside Q_2 and on the opposite side of Q_1 from the origin or outside Q_2 and on the same side of Q_1 as the origin.**

PROOF. The function Q_1 is positive (negative) according as w is on the opposite (same) side of Q_1 from (as) the origin. The function Q_2

* The case of w inside Q_2 and on the same side of Q_1 as the origin is omitted since then w would be inside Q_0.

is positive (negative) according as w is outside (inside) Q_2. Comparing these statements we have the theorem:

THEOREM 3. *A necessary and sufficient condition that* $r_e = r_i/k$, $(k \geqq 0)$, *is that* w *lie on the locus of the equation*

$$k^2(z\bar{z} - 1)^2 = \left[z^2 - (a - \bar{a})\frac{z}{c} - \frac{\bar{c}}{c}\right]\left[\bar{z}^2 + (a - \bar{a})\frac{\bar{z}}{\bar{c}} - \frac{c}{\bar{c}}\right]$$
$$= |z - \xi_1|^2 |z - \xi_2|^2,$$

where ξ_1 *and* ξ_2 *are the fixed points of the transformation* T.

PROOF. From (4), by adding $1/r_i^2$ to each member, multiplying by $r_i^2/(\nu\bar{\nu})^2$, putting $w = -\bar{\alpha}/\nu$, making use of the relation $\alpha\bar{\alpha} - \nu\bar{\nu} = 1$, putting $r_i/r_e = k$, and factoring the right-hand member, we get

$$k^2(w\bar{w} - 1)^2 = |w - \xi_1|^2 |w - \xi_2|^2, \tag{6}$$

where

$$\xi_1, \xi_2 = \{(a - \bar{a}) \pm [(a - \bar{a})^2 + 4c\bar{c}]^{1/2}\}/2c.$$

THEOREM 4. *If* T *is a hyperbolic transformation, the minimum value of* k *for all transformations* G *is given by*

$$k_0^2 = \frac{1}{4c\bar{c}}[(a - \bar{a})^2 + 4c\bar{c}] = \frac{1}{4}|\xi_1 - \xi_2|^2;$$

this value of k *is taken on if and only if* w *is on the circle*

$$Q_{3,4}(z) \equiv \frac{1}{2}(a - \bar{a})iz\bar{z} - icz + i\bar{c}\bar{z} + (a - \bar{a})\frac{i}{2}; \tag{7}$$

for parabolic transformations zero is the lower bound for k; *for elliptic transformations the minimum value of* k *is zero, and this value is taken on if and only if* $w = \xi_1$.

PROOF. Let us consider

$$k_0^2 = \frac{1}{4c\bar{c}}[(a - \bar{a})^2 + 4c\bar{c}].$$

From (6) with the right-hand member expanded in terms of the coefficients of T, we have

$$c\bar{c}(w\bar{w} - 1)^2(k^2 - k_0^2)$$
$$= (a - \bar{a})(cw - \bar{c}\bar{w})(w\bar{w} + 1) - w\bar{w}(a - \bar{a})^2 - c^2w^2 - \bar{c}^2\bar{w}^2$$
$$+ (w^2\bar{w}^2 + 1)c\bar{c} - (w\bar{w} - 1)^2(a - \bar{a})^2/4 - (w\bar{w} - 1)^2c\bar{c}$$

$$= \left[\frac{1}{2}(a-\bar{a})iw\bar{w} - icw + i\bar{c}\bar{w} + (a-\bar{a})\frac{i}{2}\right]^2$$

$$= [Q_{3,4}(w)]^2 \geqq 0$$

for all w, since the expression in brackets is real. Hence, $k^2 = k_0^2$ if and only if w is on $Q_{3,4}$ since $|w| \neq 1$ and $c \neq 0$.

Let us notice that k^2 cannot have the value k_0^2 for elliptic transformations, since in that case $(a-\bar{a})^2+4cc<0$ whereas k is a non-negative real number. Furthermore, $Q_{3,4}(w)$ has no real locus for T an elliptic transformation.

From the explicit expression for $Q_{3,4}(w)$ in the non-elliptic case it is obviously a circle with center at $-2\bar{c}/(a-\bar{a})$, a not real, and radius equal to $\{\pm i[(a-\bar{a})^2+4c\bar{c}]^{1/2}\}/(a-\bar{a})$, the sign being so chosen that the radius is positive.

Let us notice that $r_{3,4}=0$ for T parabolic, also that $Q_{3,4}(\xi)=0$, where $\xi=(a-\bar{a})/2c$ is the fixed point of the transformation. Now for parabolic transformations we have $Q_0(\xi)=0$; hence ξ is not a permissible value for w.

From (6) it is obvious that $k=0$ for elliptic transformations and that $w=\xi_1, \xi_2$. However, $w \neq \xi_2$ since $Q_0(\xi_2)<0$. We merely call attention to the fact that S is a rotation when w is equal to ξ_1, the fixed point of T which is outside Q_0.

For hyperbolic transformations ξ_1 and ξ_2 are on Q_0 and hence are not permissable values of w.

Let us now consider k_0 when a is real. From the expression given in the theorem for k_0 we see that it is unity for a real. From (6) it follows immediately that

$$(k^2-1)(w\bar{w}-1)^2c\bar{c} = -(cw-\bar{c}\bar{w})^2 \geqq 0$$

for all w since $cw-\bar{c}\bar{w}$ is a pure imaginary. Hence, for a real the minimum value of k is unity and is taken on if and only if w is on $Q_{3,4} \equiv Q_1 \equiv Q_2$. We call attention to the fact that if a is real, the transformation is hyperbolic since g_i and g_i' are on opposite ends of a diameter of Q_0.

It is readily seen from (6), with w and $\bar{w}$ as independent variables, that there are no relative maximum and minimum values of k.

REMARK. If G is so chosen that $k=k_0$, then no transform of S will increase the magnitude of its isometric circle. This is apparent since applying G_2 to $S=G_1TG_1^{-1}$ is equivalent to applying $G_3=G_2G_1$ to T, but G_1 is such that I_t receives its maximum magnification.

COROLLARY $Q_{3,4}$ *is self-inverse with respect to* Q_0.

PROOF. Replacing z and $\bar{z}$ in $Q_{3,4}$ by $1/\bar{z}$ and $1/z$, respectively, and multiplying by $z\bar{z}$, we get the equation $Q_{3,4}(z)=0$.

We found in Theorem 3 a necessary and sufficient condition that $r_0=r_1/k$ to be that w satisfy the equation.

$$k^2(z\bar{z}-1)^2=|z-\xi_1|^2|z-\xi_2|^2.$$

We shall now investigate the geometric significance of that equation. Let

$$\begin{aligned}
M &= (a-\bar{a})^2+4c\bar{c},\\
B_3 &= -\left[(a-\bar{a})+(M-4k^2c\bar{c})^{1/2}\right]/2\bar{c}(k^2-1),\\
B_4 &= -\left[(a-\bar{a})-(M-4k^2c\bar{c})^{1/2}\right]/2\bar{c}(k^2-1),\\
C_3 &= B_3/B_4=\left[a-\bar{a}+(M-4k^2c\bar{c})^{1/2}\right]^2/4c\bar{c}(k^2-1),\\
C_4 &= B_4/B_3=\left[a-\bar{a}-(M-4k^2c\bar{c})^{1/2}\right]^2/4c\bar{c}(k^2-1).
\end{aligned}$$

THEOREM 5. *A necessary and sufficient condition that* $r_0=r_1/k$, $(k\neq 1)$, *is that* w *lie on either*

$$Q_3(z)=z\bar{z}+B_3z+\overline{B}_3\bar{z}+C_3=0 \tag{8}$$

or

$$Q_4(z)=z\bar{z}+B_4z+\overline{B}_4\bar{z}+C_4=0. \tag{9}$$

PROOF. Multiplying $Q_3(z)$ by $Q_4(z)$, slightly changing the form of (6), and equating coefficients of like terms, we have the following equations that must be satisfied simultaneously:

$$B_3+B_4=-(a-\bar{a})/\bar{c}(k^2-1), \tag{10}$$

$$B_3B_4=c/\bar{c}(k^2-1), \tag{11}$$

$$B_3C_4+B_4C_3=-(a-\bar{a})/\bar{c}(k^2-1), \tag{12}$$

$$C_3C_4=1, \tag{13}$$

$$C_3+C_4+B_3\overline{B}_4+\overline{B}_3B_4=\left[(a-\bar{a})^2-2k^2c\bar{c}\right]/c\bar{c}(k^2-1). \tag{14}$$

Solving (10), (11), (12), and (13) we get the values stated for B_3, B_4, C_3, and C_4 and also $C_3=C_4=1$. It is readily seen that the given values of B_3, B_4, and $C_3=C_4=1$ do not satisfy (14). However, the given values of the four symbols do satisfy the five equations.

COROLLARY 1. *The circles* Q_3 *and* Q_4 *are always real circles for* T *an elliptic or a parabolic transformation; they are real circles for* T *a hyperbolic transformation and* $k\geq k_0$, *the minimum value of* k *for* T *hyperbolic.*

PROOF. Since C_3 and C_4 are real it is sufficient to show that the radius is real for each of the cases mentioned. The expression under the radical in B_3, B_4, C_3, and C_4 is readily seen to be non-positive for T an elliptic or a parabolic transformation and also for T hyperbolic provided $k \geqq k_0$. Therefore,

$$r_j^2 = B_j\bar{B}_j - C_j, \qquad j = 3, 4,$$
$$= -\frac{[a - \bar{a} - (M - 4k^2c\bar{c})^{1/2}]^2k^2}{4c\bar{c}(k^2 - 1)^2} \geqq 0.$$

COROLLARY 2. *Q_3 and Q_4 are inverse circles with respect to Q_0.*

PROOF. In order to find the inverse of a circle with respect to Q_0, we need only replace z and $\bar{z}$ by $1/\bar{z}$ and $1/z$, respectively, and multiply by $z\bar{z}$. The corollary is immediate upon performing this operation since $C_3C_4 = 1$, $B_3 = B_4/C_4$, and C_3 and C_4 are real.

COROLLARY 3. *In the hyperbolic case, the circles Q_3 and Q_4 coincide with*

$$Q_{3,4}(z) = \frac{1}{2}(a - \bar{a})iz\bar{z} - icz + i\bar{c}\bar{z} + (a - \bar{a})\frac{i}{2} = 0,$$

when k is equal to k_0, the minimum value of k.

PROOF. The expression under the radical in B_3, B_4, C_3, and C_4, when evaluated for $k = k_0$, is zero; hence we have $B_3 = B_4$ and $C_3 = C_4 = 1$ and therefore $Q_3 = Q_4$. Evaluating B_3 for $k = k_0$ and putting this into the equation of Q_3, along with $C_3 = 1$, we have

$$Q_3(z) = Q_4(z) = Q_{3,4}(z).$$

COROLLARY 4. *The radius r_3 of Q_3 is equal to the product of k, the reciprocal of the magnification of r_t, and the absolute value of the center $-\bar{B}_3$ of Q_3. A similar statement is true for Q_4.*

PROOF. Symbolically, the above statement is

$$r_3^2 = k^2B_3\bar{B}_3.$$

The validity of this relation is immediate upon recalling that

$$r_3^2 = B_3\bar{B}_3 - C_3,$$

equating the right-hand members of these two equations, and replacing B_3 and C_3 by their values as given immediately preceding theorem 5.

THEOREM 6. *Q_1, Q_2, Q_3, and Q_4 are fixed circles of the transformation T.*

PROOF. A necessary and sufficient condition that a given circle be a fixed circle of T is that it be orthogonal to both I_t and I_t'.*

Using the law of cosines, we find the angle θ between any two circles

$$q_j = z\bar{z} + b_j z + \bar{b}_j \bar{z} + c_j = 0, \qquad j = 1, 2,$$

to be given by

$$\cos\theta = \frac{1}{2r_1 r_2}\left[c_1 + c_2 - b_1\bar{b}_2 - \bar{b}_1 b_2\right], \tag{15}$$

$$r_j^2 = b_j\bar{b}_j - c_j, \qquad j = 1, 2.$$

Now using the analytic expressions for Q_3 and I_t and substituting in (15), we have $\cos\theta = 0$. Hence Q_3 and I_t are orthogonal. Similarly, we find that Q_3 and I_t' are orthogonal. Therefore, the circles Q_3 are fixed circles of the transformation T. It can be shown in a similar manner that Q_1, Q_2, and Q_4 are also fixed circles of T.

THEOREM 7. *For T hyperbolic*

$$k^2 = \frac{(a - \bar{a})^2 + 4c\bar{c}}{4c\bar{c}\sin^2\theta},$$

where θ is the angle between Q_0 and Q_4.

PROOF. We shall use (15), letting $q_1(z) = Q_0(z)$ and $q_2(z) = Q_4(z)$. Then

$$\cos\theta = \frac{1}{2r_4}(-1 + C_4)$$

$$= -\frac{[1 + (k^2 - 1)B_4\overline{B}_4]}{2k(B_4\overline{B}_4)^{1/2}}$$

by Corollary 4 of Theorem 5. After putting in the value of B_4 and simplifying, we have

$$\cos\theta = -\frac{[(a - \bar{a})^2 - 4c\bar{c}(k^2 - 1)]^{1/2}}{2ik(c\bar{c})^{1/2}}.$$

Now squaring and solving for k^2, we get the value stated in the theorem.

NEW MEXICO COLLEGE OF AGRICULTURE AND MECHANIC ARTS

* L. R. Ford, *Automorphic Functions*, p. 28.

A GENERALIZATION OF A PROPERTY OF HARMONIC FUNCTIONS*

MAX CORAL

1. **Introduction.** A well known theorem of Bôcher and Koebe characterizes a function $u(x, y)$ as harmonic in a region A if u is of class C' in A† and if the integral of its normal derivative is zero around every circle C in A. This theorem has been generalized by Gergen‡ as follows:

If $v(x, y)$ is harmonic and positive in A, if $u(x, y)$ is of class C' in A, and if

$$\int_C v \frac{\partial u}{\partial n} ds = \int_C u \frac{\partial v}{\partial n} ds \tag{1}$$

for every circle C interior to A, then u is harmonic in A.

The Bôcher-Koebe result is secured from Gergen's theorem by choosing $v(x, y) \equiv 1$ in A. Gergen's theorem in turn is a special case of the following theorem concerning a general linear partial differential equation of the second order which is self-adjoint and of elliptic type:

THEOREM 1. *Consider the differential expression*

$$L(z) = az_{xx} + 2bz_{xy} + cz_{yy} + dz_x + ez_y + fz \tag{2}$$

whose coefficients $a, b, \cdots, f$ are functions of (x, y) of class C'' in A which satisfy the conditions

$$a_x + b_y = d, \qquad b_x + c_y = e, \qquad ac - b^2 > 0 \tag{3}$$

in A. Let

$$\lambda(z) = az_x + bz_y + dz, \qquad \mu(z) = bz_x + cz_y + ez. \tag{4}$$

Let $v(x, y)$ be a function of class C'' which never vanishes in A and which satisfies $L(v) = 0$. If $u(x, y)$ is of class C' in A, and if

* Presented to the Society, October 28, 1933, under the title, *A property of self-adjoint elliptic partial differential equations.*

† A function $u(x, y)$ is of class $C^{(n)}$ in A if it is continuous and has continuous partial derivatives in A of all orders up to and including the nth.

‡ J. J. Gergen, *Note on a theorem of Bôcher and Koebe*, this Bulletin, vol. 37 (1931), pp. 591–596. See also S. Saks, *Note on defining properties of harmonic functions*, this Bulletin, vol. 38 (1932), pp. 380–382.

$$\int_C v[\lambda(u)dy - \mu(u)dx] = \int_C u[\lambda(v)dy - \mu(v)dx] \tag{5}$$

for every circle C in A, then $u(x, y)$ is of class C'' in A and $L(u) = 0$.

2. **The Haar fundamental lemma.** Theorem 1 is capable of rather simple proof if use is made of the so-called fundamental lemma of Haar* for the calculus of variations of a double integral. We shall establish Haar's lemma in a form adapted to our purposes:

THEOREM 2. *Let $P(x, y)$, $Q(x, y)$, and $T(x, y)$ be continuous in a bounded region A. If the integral*

$$\iint_A (P\zeta_x + Q\zeta_y + T\zeta)dx\,dy \tag{6}$$

has the value zero for every function $\zeta(x, y)$ of class C' in A which vanishes on the boundary of A, then

$$\int_C Pdy - Qdx = \iint_C Tdx\,dy \tag{7}$$

on every circle C interior to A. Conversely, if (7) holds on every circle C interior to A, and if D is a region interior to A which is bounded by a simply closed regular curve Δ, then

$$\iint_D (P\zeta_x + Q\zeta_y + T\zeta)dx\,dy = 0$$

for every function $\zeta(x, y)$ of class C' in D which vanishes on Δ.

Let C be a circle interior to A with center at (x_0, y_0) and radius r, and consider the ring Γ defined by

$$\begin{aligned} x &= x_0 + \rho\cos\theta, \\ y &= y_0 + \rho\sin\theta, \qquad 0 < r_1 \leqq \rho \leqq r;\ 0 \leqq \theta \leqq 2\pi. \end{aligned} \tag{8}$$

If $R(\rho)$, $(r_1 \leqq \rho \leqq r)$, is any function of class C' for which $R(r_1) = R(r) = 0$ and for which the derivative $R'(\rho)$ also vanishes for $\rho = r_1$ and $\rho = r$, define

$$\zeta \equiv \begin{cases} 0 & \text{on } A - \Gamma, \\ R & \text{on } \Gamma. \end{cases} \tag{9}$$

* A. Haar, *Über eine Verallgemeinerung des Du Bois-Reymondsche Lemmas*, Acta Litterarum ac Scientiarum, vol. 1 (1922), p. 33. Cf. also A. Haar, *Zur Variationsrechnung*, Abhandlungen aus dem Mathematischen Seminar der Hamburgischen Universität, vol. 8 (1930), pp. 1–27.

The function $\zeta(x, y)$ so defined is of class C' on A and vanishes on the boundary of A; and

$$\iint_A (P\zeta_x + Q\zeta_y + T\zeta)dx\,dy = \int_0^{2\pi}\int_{r_1}^{r}(PR'\cos\theta + QR'\sin\theta + TR)\rho\,d\rho\,d\theta,$$

the arguments in P, Q, and T on the right-hand side of this equality being the values (8) belonging to Γ. The hypothesis that the integral (6) vanishes for every function $\zeta(x, y)$ of class C' which vanishes on the boundary of A thus implies that

$$\int_{r_1}^{r}\left\{R'\int_0^{2\pi}(P\cos\theta + Q\sin\theta)d\theta + R\int_0^{2\pi}Td\theta\right\}\rho\,d\rho = 0$$

for every function $R(\rho)$ of the type described above. By Du Bois-Reymond's lemma* it follows that

$$\int_0^{2\pi}(P\cos\theta + Q\sin\theta)\rho\Big|_{\rho=r_1}^{\rho=r}d\theta = \int_0^{2\pi}\int_{r_1}^{r}T\rho\,d\rho\,d\theta.$$

Letting r_1 approach zero, one secures the equality (7).

Suppose, conversely, that (7) holds on every circle C interior to A. Let the minimum distance of Δ from the boundary of A be $2r>0$. If $C_{xy,\rho}$ represents a circle whose center is at a point (x, y) of the closed region $D+\Delta$ and whose radius is $\rho<r$, then the functions

$$\pi_\rho(x, y) \equiv \iint_{C_{xy,\rho}} P(x+\xi, y+\eta)d\xi\,d\eta,$$

$$\kappa_\rho(x, y) \equiv \iint_{C_{xy,\rho}} Q(x+\xi, y+\eta)d\xi\,d\eta, \quad (x, y) \text{ in } D+\Delta, \tag{10}$$

$$\tau_\rho(x, y) \equiv \iint_{C_{xy,\rho}} T(x+\xi, y+\eta)d\xi\,d\eta,$$

are of class C' in $D+\Delta$ and have derivatives which are given by

$$\pi_{\rho x} = \int_{C_{xy,\rho}} P\,d\eta, \qquad \pi_{\rho y} = -\int_{C_{xy,\rho}} P\,d\xi, \tag{11}$$

* Cf. A. Huke, *An historical and critical study of the fundamental lemma of the calculus of variations*, in *Contributions to the Calculus of Variations, 1930*, University of Chicago Press, 1931, p. 86.

with similar expressions for the derivatives of κ_ρ and τ_ρ.† Furthermore, an application of the mean value theorem shows that

$$\lim_{\rho\to 0}\frac{1}{\pi\rho^2}\pi_\rho(x, y) = P(x, y),$$

$$\lim_{\rho\to 0}\frac{1}{\pi\rho^2}\kappa_\rho(x\ \ y) = Q(x, y), \tag{12}$$

$$\lim_{\rho\to 0}\frac{1}{\pi\rho^2}\tau_\rho(x.\ y) = T(x, y),$$

uniformly for (x, y) in $D+\Delta$. Now if $\zeta(x, y)$ is any function of class C' in D which vanishes on Δ, then

$$\begin{aligned}\iint_D (P\zeta_x + Q\zeta_y + T\zeta)dx\,dy &= \lim_{\rho\to 0}\frac{1}{\pi\rho^2}\iint_D (\pi_\rho\zeta_x + \kappa_\rho\zeta_y + \tau_\rho\zeta)dx\,dy \\ &= \lim_{\rho\to 0}\frac{1}{\pi\rho^2}\iint_D [(\pi_\rho\zeta)_x + (\kappa_\rho\zeta)_y]dx\,dy \\ &\quad + \lim_{\rho\to 0}\frac{1}{\pi\rho^2}\iint_D (\tau_\rho - \pi_{\rho x} - \kappa_{\rho y})\zeta dx\,dy,\end{aligned} \tag{13}$$

as can be seen by the use of (12) and an integration by parts. The first term on the right of the last equality in (13) is zero by Green's theorem since ζ vanishes on Δ; and the second term also vanishes since

$$\pi_{\rho x}(x, y) + \kappa_{\rho y}(x, y) - \tau_\rho(x, y) = \int_{C_{xy,\rho}} (P\,d\xi - Q\,d\eta) - \iint_{C_{xy,\rho}} T\,d\xi\,d\eta$$

identically for (x, y) in $D+\Delta$ and $0<\rho<r$. Thus the second part of Theorem 2 is established.

3. **Proof of Theorem** 1. By virtue of the first two equations (3), the equation $L(z)=0$ is the Lagrange equation

$$\frac{\partial}{\partial x}F_p + \frac{\partial}{\partial y}F_q - F_z = 0 \tag{14}$$

† A proof of the analog of this property of the *integral means* π_ρ, κ_ρ, τ_ρ for the case of functions of three variables can be found in O. D. Kellogg, *Foundations of Potential Theory*, Berlin, 1929, p. 224.

for the variation problem

$$\iint_A F(x, y, z, p, q)dx\,dy = \text{minimum}$$

whose integrand is the quadratic form

$$F = \tfrac{1}{2}[ap^2 + 2bpq + cq^2 + 2dzp + 2ezq + (d_z + e_y - f)z^2],$$

where we have put $p = z_x$, $q = z_y$. Furthermore, if $F_z(z)$, $F_p(z)$, $F_q(z)$ denote the values of these derivatives with the arguments (x, y, z, z_x, z_y) belonging to z, we have

$$F_p(z) = \lambda(z), \qquad F_q(z) = \mu(z),$$

and the equation (5) becomes

$$\begin{aligned}\int_C v[F_p(u)dy - F_q(u)dx] &= \int_C u[F_p(v)dy - F_q(v)dx] \\ &= \iint_C \left\{\frac{\partial}{\partial x}[uF_p(v)] + \frac{\partial}{\partial y}[uF_q(v)]\right\}dx\,dy \\ &= \iint_C \{uF_z(v) + u_xF_p(v) + u_yF_q(v)\}dx\,dy \\ &= \iint_C \{vF_z(u) + v_xF_p(u) + v_yF_q(u)\}dx\,dy,\end{aligned} \tag{15}$$

where we have made use of Green's theorem, equation (14), and a familiar property of the quadratic form F, namely,

$$uF_z(v) + u_xF_p(v) + u_yF_q(v) = vF_z(u) + v_xF_p(u) + v_yF_q(u).$$

Let C_0 be any circle interior to A. Since (15) holds on every circle C interior to A, it follows from Theorem 2 that

$$\iint_{C_0} \{F_p(u)v\zeta_x + F_q(u)v\zeta_y + [F_z(u)v + F_p(u)v_x + F_q(u)v_y]\zeta\}dxdy \tag{16}$$

vanishes for every function $\zeta(x, y)$ of class C' in C_0 and vanishing on the boundary of C_0. Hence

$$\iint_{C_0} [F_p(u)\zeta_x^* + F_q(u)\zeta_y^* + F_z(u)\zeta^*]dx\,dy \tag{17}$$

vanishes for every function $\zeta^*(x, y)$ of class C' in C_0 which vanishes on the boundary of C_0, since for any such function ζ^* we may define

$\zeta \equiv \zeta^*/v$ and the integral (17) then reduces to the integral (16). Again using Theorem 2 we find that

$$\int_C F_p(u)dy - F_q(u)dx = \iint_C F_z(u)dx\,dy \tag{18}$$

for every circle C interior to C_0. Since C_0 is any circle in A, the result (18) holds for every circle C in A.

That the function $u(x, y)$ is of class C'' in A now follows from a result due to Lichtenstein† since $ac-b^2>0$; hence, by Green's theorem,

$$\iint_C \left[\frac{\partial}{\partial x}F_p(u) + \frac{\partial}{\partial y}F_q(u) - F_z(u)\right]dx\,dy$$

$$= \int_C [F_p(u)dy - F_q(u)dx] - \iint_C F_z(u)dx\,dy = 0$$

on every circle C interior to A. It follows readily that

$$L(u) = \frac{\partial}{\partial x}F_p(u) + \frac{\partial}{\partial y}F_q(u) - F_z(u) = 0$$

in A, as was to be proved.

WAYNE UNIVERSITY

† L. Lichtenstein, *Über den analytischen Charakter der Lösungen regulärer zweidimensionaler Variationsprobleme*, Bulletin de l'Académie des Sciences de Cracovie, A, 1912, pp. 915–941.